

Minitab®

Session Commands

Minitab®

Minitab®, Minitab Connect®, Minitab Model Ops®, Minitab Engage®, Minitab Workspace®, Minitab Simul8®, Salford Predictive Modeler®, SPM®, and the Minitab® logo are all registered trademarks of Minitab, LLC, in the United States and other countries. Additional trademarks of Minitab, LLC can be found at www.minitab.com. All other marks referenced remain the property of their respective owners.

© 2026 Minitab, LLC. All rights reserved.

Contents

Using Session Commands **11**
Alphabetical list of session commands 11
What are session commands? 28
Session command syntax notation 28
Symbols to use with session commands 29
Using a subcommand 30
Using session commands in the **Command Line** pane and the **History** pane 30
Using the **History** pane 30
Rules for entering session command arguments 31
Rules for entering session commands 32
Interrupting session command execution 32
Updates for release 19.1 32
Commands that are unavailable in the web app 38

Opening, Saving, and Printing Files **39**
END: Session command for ending data input 39
GSAVE: Session subcommand for saving a graph in a file 39
ODBC: Session command for importing data from a database file 40
OUTFILE and NOOUTFILE: Session commands for saving a Minitab session in a text file 41
PRINT: Session command for displaying columns, constants, or matrices in the output pane 42
READ data into columns 42
READ data into a matrix 44
RESTART: Session command for restarting 44
RETRIEVE: Session command for retrieving a saved worksheet or project 45
SAVE: Session command for saving a worksheet or project 45
STOP: Session command for closing Minitab 46
WOPEN: Session command for opening a worksheet 46
WORKSHEET: Session command for making a worksheet active, for closing a worksheet, or for renaming a worksheet 49
WRITE: Session command for writing data to the screen or a data file 50
WSAVE: Session command for saving a worksheet file 52
XPPOINT: Session command for sending output to Microsoft PowerPoint 54
XWORD: Session command for sending output to Microsoft Word 54

Other Session Commands **55**
PYSC: Session command for running a Python script 55
RSCR: Session command for running a R script 55
NAME: Session command for assigning names to columns, stored constants, and matrices 56
ABORT: Subcommand for exiting a multi-line command 57
HELP: Session command for opening this guide 57

Dynamic Data Exchange **58**
XDACTIVATE: Session command for activating a link 58

XDADD: Session command for adding a new link 58
XDDEACTIVATE: Session command for deactivating a client link 60
XDEXEC: Session command for executing a command in a remote application 61
XDGET: Session command for performing a one-time data transfer 61
XDREMOVE: Session command for deleting an established link 62

Manipulating and Calculating Data 64
Calculator 64
Data 107
Editor 133

Basic Statistics 140
DESCRIBE: Session command for summarizing numeric data with statistics 140
STATS: Session command for storing descriptive statistics 144
GSUMMARY: Session command for displaying a graphical summary of each variable 147
ONEZ: Session command for performing a 1-sample Z-test 148
ONET: Session command for performing a 1-sample t-test 149
TWOT: Session command for performing a 2-sample t-test when samples are in one column 150
TWOSAMPLE: Session command for performing a 2-sample t-test when the samples are in different columns. 151
PAIR: Session command for performing a paired t-test 152
PONE: Session command for performing a hypothesis test of the proportion 153
PTWO: Session command for performing a hypothesis test of the difference between two proportions 155
ONERATE: Session command for performing a 1-sample Poisson rate test 156
TWORATE: Session command for performing a 2-sample Poisson rate test 157
ONEV: Session command for performing a 1 variance test 158
TWOVARIANCES: Session command for determining whether the variances or standard deviations of two groups differ 159
CORRELATION: Session command for measuring the strength and direction of the association between two variables 161
COVARIANCE: Session command for calculating the covariance between pairs of columns 163
NORMTEST: Session command for performing a normality test 164
OUTLIER: Session command for performing an outlier test 166
PGOODNESS: Session command for performing a chi-square goodness-of-fit test for Poisson distribution 167

Regression 169
INDICATOR: Session command for creating indicator variables 169
REGRESS: Session command for performing a regression analysis 170
BREG: Session command for performing best subsets regression 179
FITLINE: Session command for creating a fitted line plot 180
SSWORKSHEET: Session command for creating a stability study worksheet 183
SHELFLIFE: Session command for performing a stability study 184
NLINEAR: Session command for performing nonlinear regression 189
OREG: Session command for performing orthogonal regression 193
PLS: Session command for performing partial least squares regression 195
GZLM: Session command for fitting a binary logistic model or a Poisson model 200
BFIT: Session command for creating a binary fitted line plot 211
OLOGISTIC: Session command for performing ordinal logistic regression 214

NLOGISTIC: Session command for performing nominal logistic regression 217

ANOVA 220
ONEWAY: Session command for performing a one-way ANOVA 220
ANOM: Session command for creating an analysis of means chart 224
ANOVA: Session command for performing a balanced ANOVA 225
GLM: Session command for fitting the general linear model 228
REML: Session command for fitting a mixed effects model 236
COMPARE: Session command for performing multiple comparisons of means 242
MANOVA: Session command for performing a general MANOVA 245
NESTED: Session command for performing a fully-nested ANOVA 249
VARTEST: Session command for performing an equal variances test 250
INTPLOT: Session command for creating an interval plot 251
MAIN: Session command for creating a main effects plot 254
INTERACT: Session command for creating an interactions plot 255

DOE 258
Screening Designs 258
Factorial Designs 279
Response Surface Designs 343
Mixture Designs 367
Taguchi Designs 414
Modify and Display Designs 421

Quality and Process Improvement 427
Quality Planning Tools 427
Control Charts 434
Capability Analysis 633
Tolerance Intervals 722
Measurement Systems Analysis 726
Acceptance Sampling 750

Reliability/Survival 757
Test Plans 757
Distribution Analysis (Right Censoring) 765
Distribution Analysis (Interval Censoring) 783
Growth Curves 800
Regression with Life Data 806
Cox Regression 811
Probit Analysis 820
Warranty Analysis 825

Predictive Analytics 828
CART Regression 828
CART Classification 835
TreeNet Regression 843
TreeNet Classification 854

Random Forests Regression 866
Random Forests Classification 871
MARS Regression 877
Discover Best Model (Continuous Response) 884
Discover Best Model (Binary Response) 897

Multivariate Analysis 911
PCA: Session command for performing principal components analysis 911
FACTOR: Session command for performing a factor analysis 912
CLUOBS: Session command for clustering observations 915
CLUVARS: Session command for clustering variables 919
KMEANS: Session command for non-hierarchical clustering of observations 921
DISCRIMINANT: Session command for performing discriminant analysis 922
ITEMANALYSIS: Session command for performing item analysis 924
CA: Session command for performing simple correspondence analysis 925
MCA: Session command for performing a multiple correspondence analysis 929

Time Series Analysis 932
AAMODEL: Session command for the selection of an alternative model from **Forecast with Best ARIMA Model** 932
ACF: Session command for calculating autocorrelation 932
ADF: Session command for conducting an augmented Dickey-Fuller test 933
ARIMA: Session command for modeling time series behavior and generating forecasts 934
ATARIMA: Session command for selecting parameters for an ARIMA model and generating forecasts 936
CCF: Session command for calculating cross correlation between two time series 941
DECOMP: Session command for performing decomposition on a time series 942
DES: Session command for performing double exponential smoothing 946
DIFFERENCES: Session command for calculating differences 950
LAG: Session command for calculating the lags of a column 951
MLAG: Session command to calculate lags of one or more columns 951
MA: Session command for calculating a moving average 952
PACF: Session command for calculating partial autocorrelation 956
SES: Session command for performing single exponential smoothing 957
TBOXCOX: Session command for performing a Box-Cox transformation on time series data 961
TREND: Session command for performing a trend analysis 962
TSWINT: Session command for performing Holt-Winters seasonal exponential smoothing 966

Tables 971
TABLE: Session command for creating one-way, two-way, and multi-way tables using categorical variables 971
TALLY: Session command for displaying a one-way table for each column 973
TCHISQUARE: Session command for performing a chi-square goodness-of-fit test 974
XTABS: Session command for displaying one-way, two-way, and multi-way tables for categorical variables 975

Nonparametric Analysis 979
1-Sample Sign 979
1-Sample Wilcoxon 979
MANN-WHITNEY: Session command for performing a Mann-Whitney test 980

FRIEDMAN: Session command for performing a Friedman test 980
KRUSKAL-WALLIS: Session command for performing a Kruskal-Wallis test 981
MOOD: Session command for performing a Mood's median test 981
RUNS: Session command for performing a runs test 981
WALSH: Session command for calculating pairwise averages 982
WDIFF: Session command for calculating pairwise differences 982
WSLOPE: Session command for calculating pairwise slopes 982

Equivalence Test 984
TOST: Session command for performing an equivalence test 984

Power and Sample Size Analysis 989
FDESIGN: Session subcommand for power and sample size for a general full factorial design 989
FFDESIGN: Session subcommand for power and sample size for a 2-level factorial design 990
ONERATE: Session subcommand for power and sample size for a 1-sample Poisson rate test 991
ONEVARIANCE: Session command for power and sample size for a 1 variance test 993
ONEWAY: Session command for power and sample size for one-way ANOVA 994
PBDESIGN: Session subcommand for power and sample size for Plackett-Burman design 995
PONE: Session subcommand for power and sample size for a 1 proportion test 996
POWER: Session command for power and sample size 997
PTWO: Session subcommand for power and sample size for a 2 proportion test 999
SSCI: Session command for estimating sample size 1000
SSTI: Session command for sample size for tolerance intervals 1001
TONE: Session command for power and sample size for a 1-sample t-test 1003
TOST: Session command for power and sample size for an equivalence test 1004
TPAIRED: Session subcommand for power and sample size for a paired t-test 1005
TTWO: Session subcommand for power and sample size for a 2-sample t-test 1006
TWORATE: Session subcommand for power and sample size for a 2-sample Poisson rate test 1008
TWOVARIANCE: Session subcommand for power and sample size for a 2 variances test 1009
ZONE: Session subcommand for power and sample size for a 1-sample Z-test 1010

Graphs 1012
CHART: Session command for creating a bar chart 1012
HMAP: Session command for creating a heatmap 1015
WAFERPLOT: Session command for creating a wafer plot 1017
PLOT: Session command for creating a scatterplot 1020
NPLOT: Session command for creating a binned scatterplot 1022
GCORRELATION: Session command for creating a correlogram 1024
MATRIXPLOT: Session command for creating a matrix of plots 1025
BUBBLEPLOT: Session command for creating a bubble plot 1028
MARGPLOT: Session command for creating a marginal plot 1030
HISTOGRAM: Session command for creating a histogram 1031
DOTPLOT: Session command for creating a dotplot 1034
STEM-AND-LEAF: Session command for creating a stem-and-leaf plot 1036
PPLOT: Session command for creating a probability plot 1036
ECDF: Session command for creating an empirical CDF plot 1041
DPLOT: Session command for creating a probability distribution plot 1043

BOXPLOT: Session command for creating a boxplot 1048
INTPLOT: Session command for creating an interval plot 1051
INDPLOT: Session command for creating an individual value plot 1053
LPLOT: Session command for creating a line plot 1056
PIECHART: Session command for creating a pie chart 1058
TSPLOT: Session command for creating a time series plot 1060
PARPLOT: Session command for creating a parallel plot 1062
ARGRAPH: Session command for creating an area graph 1063
CONTOURPLOT: Session command for creating a contour plot 1065
PLTX: Session command for creating a 3D scatterplot 1068
SURFACEPLOT: Session command for creating a surface plot 1069
Graph Options 1072

Model-Based Commands 1167
FACPLOT: Session command for creating a factorial plot 1167
MFFCUBE: Session command for creating a cube plot for fitted means 1168
MMOPT: Session command for the Response Optimizer 1169
MDIAGRAM: Session command to select multiple models for the response optimizer for commands from the Predictive Analytics module 1172
GOPT: Session command for the Response Optimizer for commands from the Predictive Analytics module 1173
MOVERCONT: Session command for creating an overlaid contour plot 1180
MSURFACE: Session command for creating a surface plot 1181
PREDICT: Session command for predicting response values 1183
MMPREDICT: Session command to calculate predictions from multiple models 1188
RMCONTOUR: Session command for creating a contour plot 1190

Macros Session Commands 1192
Structure Commands 1192
Declaration Statements 1192
Local Macro Variables 1195
Control Statements 1198
Using DOS Commands 1204
Labeling Macro Output 1204
Debugging Tools 1206
Additional Local Macro Features 1207
Commands that Affect Output 1208
Communicating with Macro Users 1211
Execs Commands 1212

Supporting Concepts 1214
Add your own function 1214
Adding comments to a macro 1214
Array table for OADESIGN 1214
Assigning attributes to groups 1215
Assigning attributes with multiple graphs and groups 1215
Bar chart functions 1216
Base position for project lines, area, and bar 1217

Box-Behnken designs 1218
Calculating a chi-square statistic for a goodness-of-fit test using session commands 1218
Calculations for FFACTORIAL 1219
Central composite designs 1219
Computing weights, or smoothed values 1221
Data requirements and formats for attribute agreement analysis 1221
Default date/time formats 1223
Designs generated by FFDESIGN 1224
Entering data for factor variables 1225
Entering data for response variables 1225
Entering patterned data for the SET session command 1226
Restricted and unrestricted mixed models 1227
Examples of entering response data for logistic regression 1227
A comparison of MSURFACE and RMCONTOUR plots 1231
Graphics options for MIXCONTOUR 1231
Graphics options for MIXOVER 1233
Graphics options for MIXSURFACE 1233
Graphics options for MSURFACE 1235
Graphics options for SIMPLEX 1237
Graphics options for SPCONT 1238
Graphics options for SPSIMP 1240
Graphs that use groups with the data display subcommands 1241
How to enter data for ANOVA and GLM 1241
How to enter data for CCHART 1242
How to enter data for GCHART 1242
How to enter data for NPCHART 1243
How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART 1244
How to enter data for TCHART 1244
How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART 1246
How to enter subgroup data 1246
How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC 1248
How to specify the model for factorial designs 1249
How to specify the model for GLM 1250
How to specify the model for response surface designs 1251
How to specify the model in ANOVA 1252
How to specify the model in MGAGE 1253
Missing values in exponential smoothing 1254
Missing values in factorial, response surface, and mixture designs 1254
Notes on subcommands that store descriptive statistics (STATS command) 1255
Numbers for colors to use in session commands 1255
Numbers for fill types to use in session commands 1256
Numbers for line types to use in session commands 1256
Numbers to use for symbols and markers in session commands 1257
Overview of DDE session commands 1258
Pearson residuals 1258
Plackett-Burman designs 1259

Prompting a user for information 1260
Reading in data from a text file 1260
Generalized linear model diagnostics and residual analysis 1260
Residual analysis and regression diagnostics 1261
Restrictions on GLM models 1262
Session commands that are not allowed in macros 1262
Simplex lattice design descriptions 1263
Time unit subcommands 1263
Using groups in graphs 1265
Using READ with FORMAT 1266
Using READ without subcommands 1266
Using the SYMBOL and COLOR subcommands for PLOT 1266
Using the POSITION and MODEL subcommands 1268
Valid format items 1269

Using Session Commands

Alphabetical list of session commands

A

AAMODEL: Session command for the selection of an alternative model from Forecast with Best ARIMA Model on page 932

AASA: Session command for acceptance sampling by attributes on page 750

ABORT: Subcommand for exiting a multi-line command on page 57

ACF: Session command for calculating autocorrelation on page 932

ACTMODEL: Session command for selecting an alternative model from Discover Best Model (Binary Response) on page 901

ACTPREDICT: Session command to predict from a model in the results of Discover Best Model (Binary Response) on page 904

ADD: Session command for addition on page 64

ADF: Session command for conducting an augmented Dickey-Fuller test on page 933

ALTTESTPLAN: Session command for creating an accelerated life test plan on page 761

ANOM: Session command for creating an analysis of means chart on page 224

ANOVA: Session command for performing a balanced ANOVA on page 225

ARDECISION: Session command for accepting or rejecting an entire lot on page 756

AREA: The session subcommand for shading the area below the data values to the base on page 1072

ARGRAPH: Session command for creating an area graph on page 1063

ARIMA: Session command for modeling time series behavior and generating forecasts on page 934

ARTMODEL: Session command for selecting an alternative model from Discover Best Model (Continuous Response) on page 888

ARTPREDICT: Session command to predict from a model in the results of Discover Best Model (Continuous Response) on page 891

ATARIMA: Session command for selecting parameters for an ARIMA model and generating forecasts on page 936

ATCAPA: Session command for performing an automated capability analysis on page 672

ATCLASS: Session command for discovering the best model type for a binary response on page 897

ATGAGE: Session command for performing an attribute agreement analysis on page 748

ATREGRESS: Session command for discovering the best model type for a continuous response on page 884

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

B

BAR: Session subcommand for representing data values with bars on page 1076

BASE: Session command for fixing a starting number for the random number generator on page 87

BBDESIGN: Session command for creating a Box-Behnken design on page 343

BCAPA: Session command for performing binomial capability analysis on page 716

BCPREDICT: Session command for predicting responses for new observations with a gradient boosted classification tree model on page 863

BCTMODEL: Session command for Hyperparameter Tuing and selecting an alternative TreeNet® Classification model from Discover Key Predictors on page 861

BCTREE: Session command for creating a gradient boosted classification tree on page 854

BFFACTORIAL: Session command for analyzing a full or fractional factorial design with a binary response on page 297

BFIT: Session command for creating a binary fitted line plot on page 211

BGFACTORIAL: Session command for analyzing a general full factorial design with a binary response. on page 313

BOXCOX: Session command for performing a Box-Cox transformation on page 434

BOXPLOT: Session command for creating a boxplot on page 1048

BRPREDICT: Session command for predicting responses for new observations for a gradient boosted regression tree model on page 851

BRTMODEL: Session command for Hyperparameter Tuning and for selecting an alternative TreeNet® Regression model from Discover Key Predictors on page 849

BRTREE: Session command for creating a gradient boosted regression tree on page 843

BTFT: Session command for calculating a 1-sample bootstrap confidence interval of a function on page 97

BTPR: Session command for calculating a 1-sample bootstrap confidence interval of a proportion on page 98

BTTM: Session command for calculating a 2-sample bootstrap confidence interval for the difference of means on page 100

BREAK: Session command for transferring control from a DO- or WHILE-loop on page 1201

BREG: Session command for performing best subsets regression on page 179

BRIEF: Session command for controlling the amount of output on page 1208

BUBBLEPLOT: Session command for creating a bubble plot on page 1028

BRSREG: Session command for analyzing a response surface design with a binary response on page 354

BSCREEN: Session command for analyzing a screening design with a binary response on page 269

BWCAPA: Session command for performing between/within capability analysis on page 636

BWCHART: Session command for creating an I-MR-R/S chart on page 449

BWSIXPAC: Session command for Between/Within Capability Sixpack on page 700

C

CA: Session command for performing simple correspondence analysis on page 925

CALL and RETURN: Session commands for passing control to another macro on page 1202

CAPA: Session command for performing a normal capability analysis on page 648

CCDESIGN: Session command for creating a central composite design on page 345

CCF: Session command for calculating cross correlation between two time series on page 941

CCHART: Session command for creating a C chart on page 571

CD: Session command for displaying or changing the current directory on page 1204

CDF: Session command for calculating the cumulative probability of an x-value on page 92

CENTER: Session command for centering data on page 81

CFAUTOMATICALLY: Session command for automatically recalculating values on page 136

CFMANUALLY: Session command for manually recalculating values on page 136

CFNOW: Session command for recalculating values now on page 137

CFORMAT: Session command for conditional formatting of worksheet cells on page 133

CHART: Session command for creating a bar chart on page 1012

CIBOX: The session subcommand for displaying a median confidence interval box on a boxplot on page 1078

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CLUOBS: Session command for clustering observations on page 915

CLUVARS: Session command for clustering variables on page 919

CMEAN: Session subcommand for connecting means with lines on a boxplot on page 1083

CMEDIAN: Session subcommand for connecting means with lines on a boxplot on page 1084

CODE: Session command for changing values in columns to new values on page 120

COFFSET and GAPWIDTH: Session subcommands for the space between clusters and items in a cluster on page 1101

COMPARE: Session command for performing multiple comparisons of means on page 242

CONCATENATE: Session command for combining text columns on page 120

CONNECT: Session subcommand for connecting points with lines on page 1085

CONTOURPLOT: Session command for creating a contour plot on page 1065

%CONTPROC: Session command for creating a contour plot

CONVERT: Session command for converting text data to numeric data, and numeric data to text data on page 122

COPY: Session command for copying data on page 131

CORRELATION: Session command for measuring the strength and direction of the association between two variables on page 161

COUNT: Session command for counting the number of values in a column on page 77

COVARIANCE: Session command for calculating the covariance between pairs of columns on page 163

COXREGRESSION: Session command for performing a Cox regression analysis on page 811

CTPREDICT: Session command for predicting responses for new observations for a classification tree on page 841

CTREE: Session command for creating a classification tree on page 835

CUSUM: Session command for creating a CUSUM chart on page 590

CUTPOINT, MIDPOINT, and NINTERVAL: Session subcommands for specifying cutpoints and midpoints on page 1087

D

DATA: Session subcommand for controlling the data region within the figure region on page 1088

DATE: Session command for changing data type to date/time on page 114

DATLAB: Session command for labeling data values on page 1089

DCAPA: Session command for performing individual distribution identification on page 633

DECOMP: Session command for performing decomposition on a time series on page 942

DEFAULT: Session command for assigning default values to subcommand arguments on page 1195

DEFINE: Session command for defining a constant matrix on page 105

DEFTEST: Session command for defining the sensitivity of the tests for special causes on page 632

DELETE: Session command for deleting rows of data on page 112

DES: Session command for performing double exponential smoothing on page 946

DESCRIBE: Session command for summarizing numeric data with statistics on page 140

DIAGONAL: Session command for creating a matrix from a column on page 105

DIFFERENCES: Session command for calculating differences on page 950

DISCRIMINANT: Session command for performing discriminant analysis on page 922

DISTRIBUTION: Session command for fitting a distribution on page 1091

DO and ENDDO: Session commands for looping through a block of commands on page 1199

DOT: Session subcommand for displaying a symbol for each data value on page 1094

DOTPLOT: Session command for creating a dotplot on page 1034

DPLOT: Session command for creating a probability distribution plot on page 1043

DSDESIGN: Session command for creating a definitive screening design on page 258

DROUND: Session command for rounding date/time values on page 123

DSET: Session command for making patterned data on page 83

DTESTPLAN: Session command for creating a demonstration test plan on page 757

DTYPE: Session command for determining the data type of a column or a constant on page 1197

E

ECDF: Session command for creating an empirical CDF plot on page 1041

EIGEN: Session command for calculating eigenvalues on page 64

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

ELSE, ELSEIF, IF, ENDIF: Session commands for executing code depending on a logical condition on page 1198

END: Session command for ending data input on page 39

ENDLAYOUT and LAYOUT: Session subcommands for specifying where a graph appears on a page on page 1116

ENDMACRO, GMACRO, and MACRO: Session commands for marking the beginning and ending of a macro on page 1192

ENDWHILE and WHILE: Session commands for repeating a block of commands depending on a logical expression on page 1199

ERASE: Session command for erasing variables on page 114

ETESTPLAN: Session command for creating an estimation test plan on page 759

EVDESIGN: Session command for creating an extreme vertices design on page 367

EWMACHART: Session command for creating an EWMA chart on page 583

EXCLUDE and INCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

EXECUTE: Session command for running an Exec file on page 1213

EXIT: Session command for transferring control back to Minitab or for closing Minitab on page 1203

F

FACPLOT: Session command for creating a factorial plot on page 277

FACTOR: Session command for performing a factor analysis on page 912

FDATE/TIME: Session command for changing the format of date/time columns on page 137

FDESIGN: Session command for creating a general full factorial design on page 279

FDESIGN: Session subcommand for power and sample size for a general full factorial design on page 989

FFACTORIAL: Session command for analyzing a full or fractional factorial design on page 289

FFCUBE: Session command for creating a cube plot on page 341

FFDESIGN: Session command for creating a full or fractional factorial design on page 280

FFDESIGN: Session subcommand for power and sample size for a 2-level factorial design on page 990

FFINT: Superseded by FACPLOT except for mixture designs on page 337

FFMAIN: Superseded by FACPLOT except for mixture designs on page 339

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

FISHBONE: Session command for creating a cause-and-effect diagram on page 427

FITD: Session command for fitting a distribution to the data on a probability plot on page 1097

FITLINE: Session command for creating a fitted line plot on page 180

FNUMERIC: Session command for changing columns to numeric format on page 137

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

%FORM: Session command for creating a data collection form for a 3-factor design

FORMULA: Session command for assigning a formula to a column on page 138

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

FRIEDMAN: Session command for performing a Friedman test on page 980

FTEXT: Session command for changing the format of text columns on page 138

G

GAGERR: Session command for performing a crossed gage R&R study on page 736

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

GAWORKSHEET: Session command for creating a gage R&R study worksheet on page 728

GCHART: Session command for creating a G chart on page 622

GCORRELATION: Session command for creating a correlogram on page 1024

GENVAR: Session command for creating a generalized variance chart on page 612

GFACTORIAL: Session command for fitting a general full factorial design on page 306

GLM: Session command for fitting the general linear model on page 228

GMACRO, MACRO, and ENDMACRO: Session commands for marking the beginning and ending of a macro on page 1192

GOTO and MLABEL: Session commands for branching to any line in a macro on page 1201

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

GSAVE: Session subcommand for saving a graph in a file on page 39

GSCALE: Session command to determine appropriate scaling for a graph on page 1210

GSUMMARY: Session command for displaying a graphical summary of each variable on page 147

GZLM: Session command for fitting a binary logistic model or a Poisson model on page 200

H

HELP: Session command for opening this guide on page 57

HISTOGRAM: Session command for creating a histogram on page 1031

HLABEL: Session subcommand for labeling histogram bars on a marginal plot with y-axis values on page 1103

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

HMAP: Session command for creating a heatmap on page 1015

I

ICHART: Session command for creating an I chart on page 495

IDIDENTIFICATION: Session command for creating probability plots of interval-censored, failure (or survival) data on page 783

IDOVIEW: Session command for creating a layout of distribution plots of interval-censored data on page 786

IECTREE: Session command for selecting an alternative classification tree on page 840

IERTREE: Session command for selecting an alternative regression tree on page 833

IF, ELSEIF, ELSE, ENDIF: Session commands for executing code depending on a logical condition on page 1198

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot on page 1106

IMRCHART: Session command for creating an I-MR chart on page 483

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

INDICATOR: Session command for creating indicator variables on page 86

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot on page 1109

INDPLOT: Session command for creating an individual value plot on page 1053

INFO: Session command for summarizing the current worksheet on page 107

INTBAR: Session subcommand for displaying a vertical line with horizontal lines at the endpoints of the confidence interval for the mean on page 1110

INTERACT: Session command for creating an interactions plot on page 255

INTLAB: Session subcommand for labeling interval bar endpoints on an interval plot on page 1112

INTPLOT: Session command for creating an interval plot on page 251

INVCDF: Session command for calculating the variable for a cumulative probability on page 95

INVERT: Session command for replacing a matrix value with its inverse on page 106

IQRBOX: Session subcommand for displaying an interquartile range box on a boxplot on page 1114

ITEMANALYSIS: Session command for performing item analysis on page 924

J

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

JOHNSON: Session command for applying the Johnson transformation on page 636

K

KKCAT, KKNAME, and KKSET: Session commands for using text on page 1195

KMEANS: Session command for non-hierarchical clustering of observations on page 921

KRUSKAL-WALLIS: Session command for performing a Kruskal-Wallis test on page 981

L

LAG: Session command for calculating the lags of a column on page 64

LAYOUT and ENDLAYOUT: Session subcommands for specifying where a graph appears on a page on page 1116

LBRIGHT: Session subcommand for specifying the brightness of the lights that illuminate a surface plot on page 1117

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

LET: Session command for correcting a number in a worksheet or performing arithmetic on page 64

LIGHT: Session subcommand for specifying the position, color, and visibility of the lights that illuminate a surface plot on page 1120

LINE: The session subcommand for constructing a line from points on a graph on page 1121

LNGAGE: Session command for performing a gage linearity and bias study on page 734

LONGMETHOD: Session command for performing an attribute gage study (analytic method) on page 746

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot on page 1122

LPLOT: Session command for creating a line plot on page 1056

LREGRESSION: Session command for performing a regression analysis when the error distribution is Weibull, smallest extreme value, exponential, log-normal, normal, logistic, or log-logistic on page 806

LTABLE: Session command for fitting a distribution to interval-censored data or warranty data on page 789

LTEST: Session command for parametric or nonparametric distribution analysis on page 771

M

MA: Session command for calculating a moving average on page 952

MACHART: Session command for creating a moving average chart on page 576

MACRO, ENDMACRO, and GMACRO: Session commands for marking the beginning and ending of a macro on page 1192

MAIN: Session command for creating a main effects plot on page 254

MANN-WHITNEY: Session command for performing a Mann-Whitney test on page 980

MANOVA: Session command for performing a general MANOVA on page 245

MARGPLOT: Session command for creating a marginal plot on page 1030

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

MARS: Session command for creating a multi-adaptive regression spline model on page 877

MATRIXPLOT: Session command for creating a matrix of plots on page 1025

MCA: Session command for performing a multiple correspondence analysis on page 929

MCAPA: Session command for performing a normal capability analysis for multiple variables on page 679

MCONSTANT, MCOLUMN, MMATRIX, and MTYPE: Session commands for declaring variables on page 1192

MDESIGN: Session command to modify the design properties or display of a design in the worksheet on page 421

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot on page 1124

MEAN: Session command for calculating the arithmetic mean of a column on page 77

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot on page 1126

MEDIAN: Session command for identifying the median of a column on page 77

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot on page 1127

MEDLAB: Session subcommand for labeling medians on a boxplot on page 1128

MERGE: Session command for merging two worksheets into one worksheet on page 107

MESH: Session command for making mesh data on page 85

MEWMA: Session command for creating a multivariate EWMA chart on page 617

MFFCUBE: Session command for creating a cube plot for fitted means on page 337

MFREE: Session command for declaring a free variable on page 1193

MGAGE: Session command for performing an expanded gage R&R study on page 730

MGRID, GRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

MIDPOINT, CUTPOINT, and NINTERVAL: Session subcommands for specifying cutpoints and midpoints on page 1087

MIXCONTOUR: Session command for creating a contour plot on page 393

MIXOVER: Session command for creating an overlaid contour plot on page 401

MIXREG: Session command for analyzing a mixture design on page 377

MIXSURFACE: Session command for creating a surface plot on page 398

MLABEL and GOTO: Session commands for branching to any line in a macro on page 1201

MLAG: Session command to calculate lags of one or more columns on page 951

MMATRIX, MCONSTANT, MCOLUMN, and MTYPE: Session commands for declaring variables on page 1192

MMOPT: Session command for the Response Optimizer on page 1169

MMPREDICT: Session command to calculate predictions from multiple models on page 1188

MNCAPA: Session command for performing nonnormal capability analysis for multiple variables on page 693

MOOD: Session command for performing a Mood's median test on page 981

MOVERCONT: Session command for creating an overlaid contour plot on page 1180

MRCHART: Session command for creating an MR chart on page 502

MROPT: Session command for the Response Optimizer on page 411

MRPREDICT: Session command to predict from a multiple regression model that was an alternative model from Discover Best Model (Continuous Response) on page 894

MSURFACE: Session command for creating a surface plot on page 1181

MSPREDICT: Session command to predict from a MARS® model on page 882

MTITLE: Session command for adding a title above output on page 1204

MTYPE, MMATRIX, MCONSTANT, and MCOLUMN: Session commands for declaring variables on page 1192

MULTIPLY: Session command for multiplication on page 76

MVARCHART: Session command for creating a multi-vari chart on page 429

N

N: Session command for counting the nonmissing values in a column on page 77

NAME: Session command for assigning names to columns, stored constants, and matrices on page 56

NESTED: Session command for performing a fully-nested ANOVA on page 249

NEXT: Session command for transferring control from a loop to the beginning of the block on page 1200

NGROWTH: Session command for performing a nonparametric analysis using a nonparametric growth curve on page 803

NINTERVAL, CUTPOINT, and MIDPOINT: Session subcommands for specifying cutpoints and midpoints on page 1087

NLINEAR: Session command for performing nonlinear regression on page 189

NLOGISTIC: Session command for performing nominal logistic regression on page 217

NMISS: Session command for counting the missing values in a column on page 77

NMVARCHART: Session command for creating variability charts on page 429

NNCAPA: Session command for performing a nonnormal capability analysis on page 661

NNSIXPACK: Session command for Nonnormal Capability Sixpack on page 711

NNTINTERVALS: Session command for calculating tolerance intervals on page 724

NOBRUSH: Session subcommand for disabling brushing on a graph on page 1210

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOECHO and ECHO: Session commands for displaying Minitab commands in the output on page 1206

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

NOGRID, GRID, MGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

NOJITTER and JITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

NOLEGEND and LEGEND Session subcommands for controlling the legend on a graph on page 1118

NOMGRID GRID, MGRID, and NOGRID, and: Session subcommands for controlling the grid on a graph on page 1102

NOMISS and NOEMPTY: Session subcommands for excluding missing data from graphs on page 1130

NOOUTFILE and OUTFILE: Session commands for saving a Minitab session in a text file on page 41

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOPROPORTIONAL and PROPORTIONAL: Session subcommands for making the boxes on a boxplot proportional to the square root of the number of observations in the boxes on page 1139

NORMTEST: Session command for performing a normality test on page 164

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

NOTABLE and TABLE: Session subcommand for controlling the table within the figure region on page 1158

NOTE: Session command for adding comments that are displayed in the output on page 1211

NOTRANSPOSE and TRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

NPCAPA: Session command for performing a nonparametric capability analysis on page 668

NPCHART: Session command for creating a chart for the number of defectives on page 554

NTGAGE: Session command for performing a nested gage R&R study on page 742

NUMERIC: Session command for changing the data format of a date/time column or extracting date/time components on page 124

O

OADESIGN: Session command for creating a Taguchi orthogonal array design on page 414

OAPREDICT: Session command for calculating predicted response values on page 420

ODBC: Session command for importing data from a database file on page 40

OLAB: Session subcommand for labeling outliers on a boxplot on page 1131

OLOGISTIC: Session command for performing ordinal logistic regression on page 214

ONERATE: Session command for performing a 1-sample Poisson rate test on page 156

ONERATE: Session subcommand for power and sample size for a 1-sample Poisson rate test on page 991

ONET: Session command for performing a 1-sample t-test on page 149

ONEV: Session command for performing a 1 variance test on page 158

ONEVARIANCE: Session command for power and sample size for a 1 variance test on page 993

ONEWAY: Session command for power and sample size for one-way ANOVA on page 994

ONEWAY: Session command for power and sample size for one-way ANOVA on page 994

ONEZ: Session command for performing a 1-sample Z-test on page 148

OPTDES: Session command for selecting an optimal design on page 363

OREG: Session command for performing orthogonal regression on page 193

OUTFILE and NOOUTFILE: Session commands for saving a Minitab session in a text file on page 41

OUTLIER: Session command for performing an outlier test on page 166

OUTLIER: Session subcommand for displaying a symbol for each outlier on a boxplot on page 1133

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

P

PACF: Session command for calculating partial autocorrelation on page 956

PAIR: Session command for performing a paired t-test on page 152

PANEL: Session subcommand for paneling graphs on page 1134

PARETO: Session command for creating a Pareto chart on page 430

PAUSE and RESUME: Session commands for pausing and resuming a macro on page 1204

PBDESIGN: Session command for creating a Plackett-Burman design on page 259

PBDESIGN: Session subcommand for power and sample size for Plackett-Burman design on page 995

PCA: Session command for performing principal components analysis on page 911

PCAPA: Session command for performing a Poisson capability analysis on page 719

PCHART: Session command for creating a P chart on page 544

PDF: Session command for calculating the probability distribution of a continuous random variable on page 90

PDIAGNOSTIC: Session command for determining whether to use a P chart or a Laney P' chart on page 543

PGOODNESS: Session command for performing a chi-square goodness-of-fit test for Poisson distribution on page 167

PGROWTH: Session command for performing a parametric analysis using a parametric growth curve on page 800

PIECHART: Session command for creating a pie chart on page 1058

PLOT: Session command for creating a scatterplot on page 1020

PLS: Session command for performing partial least squares regression on page 195

PLTX: Session command for creating a 3D scatterplot on page 1068

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

PONE: Session command for performing a hypothesis test of the proportion on page 153

PONE: Session subcommand for power and sample size for a 1 proportion test on page 996

POWER: Session command for power and sample size on page 997

PPLOT: Session command for creating a probability plot on page 1036

PPRIMECHART: Session command for creating a Laney P' chart on page 549

PREDICT: Session command for predicting response values on page 1183

PRINT: Session command for displaying columns, constants, or matrices in the output pane on page 42

PROBIT: Session command for performing probit analysis on page 820

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

PROPORTIONAL and NOPROPORTIONAL: Session subcommands for making the boxes on a boxplot proportional to the square root of the number of observations in the boxes on page 1139

PTILES: Session subcommand for specifying the location of the percentile lines on a graph on page 1139

PTWO: Session command for performing a hypothesis test of the difference between two proportions on page 155

PTWO: Session subcommand for power and sample size for a 2 proportion test on page 999

PYSC: Session command for running a Python script on page 55

R

RANDOM: Session command for generating random data on page 88

RANGE: Session command for calculating a range of values in a column on page 78

RANK: Session command for ranking values in a column on page 111

RBOX: Session subcommand for displaying a range box on a boxplot on page 1143

RCHART: Session command for creating an R chart on page 464

RCOUNT: Session command for counting missing and nonmissing values in a row on page 78

RCFOREST: Session command for creating a random forest of classification trees on page 871

RCPREDICT: Session command for predicting responses for new observations with a random classification forest model on page 875

RDIDENTIFICATION: Session command for creating a distribution ID plot on page 765

RDOVIEW: Session command for creating a layout of distribution plots on page 768

READ data into columns on page 42

READ data into a matrix on page 44

READ, TSET, and SET: Session command for asking users questions and using the answers in a macro on page 1211

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

REGRESS: Session command for performing a regression analysis on page 170

REGRESS: Session subcommand for fitting a regression line to data on a graph on page 1148

REML: Session command for fitting a mixed effects model on page 236

RESTART: Session command for restarting on page 44

RESUME and PAUSE: Session commands for pausing and resuming a macro on page 1204

RETRIEVE: Session command for retrieving a saved worksheet or project on page 45

RETURN and CALL: Session commands for passing control to another macro on page 1202

RFORMULA: Session command for removing formulas on page 138

RMAXIMUM: Session command for identifying the maximum value in each row on page 78

RMCONTOUR: Session command for creating a contour plot on page 1190

RMEAN: Session command for calculating the arithmetic mean in each row on page 79

RMEDIAN: Session command for identifying the median in each row on page 79

RMERGE: Session command for merging worksheets on page 132

RMINIMUM: Session command for identifying the minimum value in each row on page 79

RN: The session command for counting the nonmissing values in a row on page 79

RNGAGE: Session command for creating gage run chart on page 729

RNMISS: Session command for counting the missing values in a row on page 79

RNMN: Session command for performing a 1-sample randomization test of a mean on page 101

RNPR: Session command for performing a 1-sample randomization test of a proportion on page 102

RNTM: Session command for performing a 2-sample randomization test of means on page 103

ROBUST: Session command for analyzing a Taguchi design on page 416

ROWTOC: Session command for stacking multiple columns into one column on page 114

RRANGE: Session command for calculating the range in each row on page 80

RRFOREST: Session command for creating a random forest of regression trees on page 866

RRPREDICT: Session command for predicting responses for new observations for a random forest regression model on page 870

RSCR: Session command for running a R script on page 55

RSREG: Session command for analyzing a response surface design with least squares regression on page 347

RSSQ: Session command for calculating the uncorrected sum of squares on page 80

RSTDEV: Session command for calculating the standard deviation in each row on page 80

RSUM: Session command for adding the values in each row on page 80

RTPREDICT: Session command for predicting responses for new observations for a regression tree on page 833

RTREE: Session command for creating a regression tree on page 828

RUNCHART: Session command for creating a run chart on page 432

RUNS: Session command for performing a runs test on page 981

S

SAMARS: Session command for selecting an alternative MARS model with a specific number of basis functions on page 882

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SAMPLE: Session command for generating rows of random data from specified columns on page 90

SAVE: Session command for saving a worksheet or project on page 45

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

SCDESIGN: Session command for creating a simplex centroid design on page 371

SCHART: Session command for creating an S chart on page 471

SCREEN: Session command for analyzing a screening design on page 261

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

SES: Session command for performing single exponential smoothing on page 957

SET: Session command for entering data into a column on page 81

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot on page 1152

SHELFLIFE: Session command for performing a stability study on page 184

SIMPLEX: Session command for creating a simplex design plot on page 385

%SIMPROC: Session command for creating a simplex plot on page 390

SINTERVAL: Session command for calculating a sign confidence interval on page 979

SIXPACK: Session command for Normal Capability Sixpack on page 705

SLABEL: Session subcommand for labeling the slices on a pie chart on page 1153

SLDESIGN: Session command for creating a mixture design on page 374

SLICE: Session subcommand for specifying the attributes of slices on a pie chart on page 1154

SOFFSET: Session subcommand for offsetting points from the center on page 1155

SORT: Session command for sorting columns on page 109

SPCONT: Session command for creating a contour design plot on page 395

SPDESIGN: Session command for creating a 2-level split-plot design on page 285

SPFACTORIAL: Session command for analyzing a 2-level split-plot design on page 322

SPLIT: Session command for splitting a worksheet into multiple worksheets on page 113

SPSIMP: Session command for creating a simplex design plot on page 387

SSCI: Session command for estimating sample size on page 1000

SSQ: Session command for calculating the uncorrected sum of squares on page 78

SSTI: Session command for sample size for tolerance intervals on page 1001

SSWORKSHEET: Session command for creating a stability study worksheet on page 183

STACK: Session command for stacking blocks of columns and constants on top of each other on page 127

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

STATS: Session command for storing descriptive statistics on page 144

STDEV: Session command for calculating the standard deviation of all the values in a column on page 78

STEM-AND-LEAF: Session command for creating a stem-and-leaf plot on page 1036

STEST: Session command for performing a 1-sample sign test on page 979

STOP: Session command for closing Minitab on page 46

SUBSET: Session command for copying specified rows to a new worksheet on page 115

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

SUM: Session command for adding the values in a column on page 78

SURFACEPLOT: Session command for creating a surface plot on page 1069

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

SYMPLOT: Session command for creating a symmetry plot on page 434

T

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

TABLE: Session command for creating one-way, two-way, and multi-way tables using categorical variables on page 971

TALLY: Session command for displaying a one-way table for each column on page 973

TBOXCOX: Session command for performing a Box-Cox transformation on time series data on page 961

TCHART: Session command for creating a T chart on page 627

TCHISQUARE: Session command for performing a chi-square goodness-of-fit test on page 974

TEXT: Session command for changing the data type of a column to text on page 125

TEXT: Session subcommand for displaying text on a graph on page 1161

TITLE: Session subcommand for adding a title to a graph on page 1162

TOGAGE: Session command for performing a Type 1 gage study on page 726

TOLINTERVALS: Session command for calculating tolerance intervals on page 722

TONE: Session command for power and sample size for a 1-sample t-test on page 1003

TOST: Session command for performing an equivalence test on page 984

TOST: Session command for power and sample size for an equivalence test on page 1004

TPAIRED: Session subcommand for power and sample size for a paired t-test on page 1005

TRACE: Session command for creating a response trace plot on page 391

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

TRANSPOSE: Session command for changing rows to columns, and columns to rows on page 106

TREND: Session command for performing a trend analysis on page 962

TSET: Session command for creating data that follow complicated patterns on page 82

TSGV: Session command for creating a Tsquared-generalized variance chart on page 602

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

TSPLOT: Session command for creating a time series plot on page 1060

TSQUARED: Session command for creating a Tsquared chart on page 607

TWARRANTY: Session command for pre-processing warranty data on page 825

TSWINT: Session command for performing Holt-Winters seasonal exponential smoothing on page 966

TTWO: Session subcommand for power and sample size for a 2-sample t-test on page 1006

TWORATE: Session command for performing a 2-sample Poisson rate test on page 157

TWORATE: Session subcommand for power and sample size for a 2-sample Poisson rate test on page 1008

TWOSAMPLE: Session command for performing a 2-sample t-test when the samples are in different columns on page 151

TWOT: Session command for performing a 2-sample t-test when samples are in one column on page 150

TWOVARIANCES: Session command for determining whether the variances or standard deviations of two groups differ on page 159

TWOVARIANCE: Session subcommand for power and sample size for a 2 variances test on page 1009

U

UCHART: Session command for creating a U chart on page 565

UDIAGNOSTIC: Session command for determining whether to use a U chart or a Laney U' chart on page 570

UNSTACK: Session command for separating a column into multiple columns on page 128

UPRIMECHART: Session command for creating a Laney U' chart on page 559

V

VARTEST: Session command for performing an equal variances test on page 250

VASAMPLING: Session command for creating or comparing variables acceptance plans on page 753

VASPECT: Session subcommand for specifying the aspect ratio of the data box on page 1164

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph on page 1165

VFACTORIAL: Session command for analyzing variability in a 2-level factorial design on page 328

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units on page 1165

VMASK: Session command for creating a two-sided CUSUM chart on page 596

VORDER: Session command for controlling the order for text categories to be processed by Minitab commands on page 138

VPOSITION: Session subcommand for specifying the view position as a ratio on page 1165

VPREPROCESS: Session command to preprocess responses for analyze variability on page 327

VUP: Session subcommand for specifying which direction is up in relation to the data box on page 1165

W

WAFERPLOT: Session command for creating a wafer plot on page 1017

WALSH: Session command for calculating pairwise averages on page 982

WDIFF: Session command for calculating pairwise differences on page 982

WHEELER: Session command for performing a crossed Evaluate Measurement Process (EMP) study on page 740

WHILE and ENDWHILE: Session commands for repeating a block of commands depending on a logical expression on page 1199

WHISKER: Session subcommand for controlling the display of whiskers on a boxplot on page 1166

WINTERVAL: Session command for calculating a Wilcoxon confidence interval on page 980

WOPEN: Session command for opening a worksheet on page 46

WORKSHEET: Session command for making a worksheet active, for closing a worksheet, or for renaming a worksheet on page 49

WPREDICTIONS: Session command for performing warranty predictions on page 825

WRITE: Session command for writing data to the screen or a data file on page 50

WSAVE: Session command for saving a worksheet file on page 52

WSLOPE: Session command for calculating pairwise slopes on page 982

WSTACK: Session command for stacking worksheets on page 129

WTEST: Session command for performing a 1-sample Wilcoxon test on page 979

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

X

XBARCHART: Session command for creating an Xbar chart on page 457

XDACTIVATE: Session command for activating a link on page 58

XDADD: Session command for adding a new link on page 58

XDDEACTIVATE: Session command for deactivating a client link on page 60

XDEXEC: Session command for executing a command in a remote application on page 61

XDGET: Session command for performing a one-time data transfer on page 61

XDREMOVE: Session command for deleting an established link on page 62

XPPOINT: Session command for sending output to Microsoft PowerPoint on page 54

XRCHART: Session command for creating an Xbar-R chart on page 435

XSCHART: Session command for creating an Xbar-S chart on page 442

XTABS: Session command for displaying one-way, two-way, and multi-way tables for categorical variables on page 975

XWORD: Session command for sending output to Microsoft Word on page 54

Z

ZMRCHART: Session command for creating a Z-MR chart on page 490

ZONE: Session command for creating a zone chart on page 477

ZONE: Session subcommand for power and sample size for a 1-sample Z-test on page 1010

What are session commands?

Session commands are a command language that you use, instead of the menus and interface, to access most functions in Minitab. Session commands are especially useful in macros.

You can type session commands in the **Command Line** pane.

Session command syntax notation

Generally, the syntax of a command includes the command name, and then one or more arguments upon which the command operates.

Usually a command operates on one or more arguments (also called parameters) that you specify; they can be columns, constants, matrices, numbers, file names, or text strings. For example, "C1" is the argument in the following command, which tells Minitab to draw a histogram of the data in column C1:

```
HISTOGRAM C1
```

Minitab Help uses the following typographical conventions for describing the syntax of individual commands.

K

Denotes a constant such as 8.3 or K14.

C

Denotes a column such as C13 or 'Height'.

C...C

Denotes a list of one or more columns separated by spaces.

M

Denotes a matrix such as M5.

E

Denotes either a constant or column, and sometimes a matrix.

[ ]

Denotes an optional argument, for example [K1].

T

Denotes a model from the project. These values are case-sensitive.
The following table provides the valid model types.

T	Corresponding Analysis
model_diagram	Response Optimizer from the Predictive Analytics Module
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)
classification_treenet	TreeNet® Classification
classification_random_forest	Random Forests® Classification
classification_tree	CART® Classification
regression_treenet	TreeNet® Regression
regression_random_forest	Random Forests® Regression
regression_tree	CART® Regression
mars_regression	MARS® Regression

Symbols to use with session commands

Use the following symbols with any session commands or subcommands.

Comment symbol (#)

Place the comment symbol # anywhere on the line to tell Minitab to ignore the rest of the line. For example:

```
DESCRIBE C1 #This is a comment
```

Missing value symbol (*)

Place the missing values symbol * anywhere a number would normally interact, to represent values that could be missing. The asterisk should be enclosed in single quotation marks ('*'). You could use the following command to copy data from one column to another, omitting rows that have missing values:

```
COPY C1 C2;
OMIT C1 = '*'.
```

Using a subcommand

Many session commands have subcommands. To use a subcommand, complete the following steps.

Note When you use a command with no subcommands, you do not have to type any punctuation mark after the command line.

1. Type the main command and end the main command line with a semicolon (;).
2. Press **Enter** to move to the next line.
3. Type as many subcommands as you need, ending each with a semicolon (;) and pressing **Enter** after each.
4. End the last subcommand with a period (.).
   If you forget to end the last subcommand with a period, you can type the period all by itself on the next line.

Using session commands in the **Command Line** pane and the **History** pane

Sometimes, it is convenient to copy a previously executed session command (or sequence of session commands) from the **Command Line** pane or the **History** pane, make minor changes if necessary, then execute the changed command(s).

Use the **Command Line** pane to type, edit, and enter commands.

The **Command Line** pane and the **History** pane can be viewed together, docked on the right side of the application frame. If the **Command Line** pane and the **History** pane are not visible, then do one of the following:

- Choose **View** > **Command Line/History**.
- Use Ctrl+K shortcut to open the **Command Line**/**History** panes.

  Note If the panes are already in view, Ctrl+K puts focus into the **Command Line** pane, but will not toggle the pane closed.

Executing session commands in the **Command Line** pane

There are several ways to enter command language into the **Command Line** pane. For instance, you can:

- Type the commands and subcommands directly into the pane. Use the **Enter** key to go to the next line. For information on arguments, go to Session command syntax notation on page 28 or Using a subcommand on page 30.
- Highlight the commands in the **History** pane and either copy and paste to the **Command Line** pane, or click **Copy to Command Line**.
- Paste text from other applications.

Press **Run** to execute the command language.

Using the **History** pane

The **History** pane provides a convenient list of the commands that you have used in your project.

You can select commands and subcommands as list items and can multi select both contiguous and non-contiguous items in the list. Selected items can be copied to clipboard or dragged to the **Command Line** pane. When pasted into the **Command Line** pane, you can edit the command language before running it.

You can also create a Minitab macro file for a routine analysis by first working through the steps using the menus, highlighting portions of session commands, and saving them as a macro or text file type.

Print the contents of the **History** pane

You can print the contents of the **History** pane. Right-click the selected text, then choose **Print History**.

Save the contents of the **History** pane

You can save the contents of the **History** pane in a text file. Right-click the selected text, then choose **Save History As**. Choose your file name and file type.

Rules for entering session command arguments

Arguments specify data characteristics, such as location or titles. They can be variables (columns, constants, matrices) as well as text strings or numbers. For information on argument notation, go to Session command syntax notation on page 28.

Variables

- Enclose variable names in single quotation marks (for example, HISTOGRAM 'Salary'). Certain commands, such as ANOVA, GLM, and the high-resolution graphics commands do not require quotation marks, but all commands work properly when quotes are used.
- In arguments, variable names and variable numbers can be used interchangeably. For example, the two following commands do the same thing (if C1 is named 'Sales'):

  ```
  DESCRIBE C1 C2
  DESCRIBE 'Sales' C2
  ```

- You can abbreviate a consecutive range of columns, stored constants, or matrices with a dash. For example, PRINT C2-C5 is equivalent to PRINT C2 C3 C4 C5.
- You can use a stored constant (such as K20) in place of any constant. You can even use stored constants to form a range such as K20:15, which represents all integers from the value of K20 to 15.

Text strings

Enclose text strings, such as labels or file names, in double quotes (for example, TITLE "This is My Title"). In earlier versions of Minitab, text was enclosed in single quotes. Although this still works, it is no longer recommended, and can cause a conflict with column and constant names.

Numbers

- Do not enclose numbers in quotes unless you want the numbers to appear as text.
- To specify a range of numbers, abbreviate the sequence using these conventions:

  1:4 expands to 1 2 3 4

  4:1 expands to 4 3 2 1

  1:3/.5 expands to 1.0 1.5 2.0 2.5 3.0

The session command SET on page 81 includes additional abbreviation conventions.

Rules for entering session commands

A session command consists of one main command, and may have one or more subcommands. Arguments and symbols may also be included in the command.

- Subcommands, which further define how the main command should be carried out, are optional unless otherwise specified.
- Arguments, which specify data characteristics, may be included one or more times for both the main command and subcommands.
- Symbols, which assist in controlling the session language, can also be included in session commands.
- Commands and column names are not case-sensitive. You can type them in lowercase, uppercase, or any combination.
- You can abbreviate any session command or subcommand by using the first four letters.
- Enter only one command or subcommand per line.
- Close all but the last line of a command with semicolons. Close the last line with a period.
- Some subcommands have their own subcommands. The order in which you give these subcommands determines what subcommand or command they modify. You can use many subcommands more than once in a command.

Note Some commands, called %macros, are macros that are invoked by typing % followed by the full macro name (you cannot abbreviate macro names).

Interrupting session command execution

To interrupt the display of data from a command or the execution of a macro, type **Ctrl**+**Break**. In a macro, Minitab finishes executing the current command, then exits the macro. Display of data is halted as soon as possible.

Updates for release 19.1

The following section includes the session commands that are new, changed or obsolete for Minitab 19.

Input and history

In Minitab 19.1, choose **View** > **Command Line/History** to open the **Command Line** pane and the **History** pane. In the **Command Line** pane, enter session commands and run macros. Use the **History** pane to see session commands that ran and to copy those session commands.

Obsolete continuation character (&)

In previous versions the & symbol indicated that a command continued on the next line, for example:

```
PLS C18 = C1-C17 c1*c2 c1*c3 c1*c4 c1*c5 c1*c6 c1*c7 c1*c8 c1*c9 c1*c10 c1*c11&
 c1*c12 c1*c13 c1*c14 c1*c15 c1*c16 c1*c17;
```

In Minitab 19.1, session commands with an & symbol create errors. Instead, type everything on 1 line.

```
PLS C18 = C1-C17 c1*c2 c1*c3 c1*c4 c1*c5 c1*c6 c1*c7 c1*c8 c1*c9 c1*c10 c1*c11 c1*c12
 c1*c13 c1*c14 c1*c15 c1*c16 c1*c17;
```

Adding comments and notes

In Minitab 19.1, each instance of the NOTE command creates a new output tab. To keep different notes together, surround all of the notes and output that you want on one output tab with MTITLE and ENDMTITLE.

New commands

Resampling commands

BTFT
Bootstrapping for 1-sample function

BTPR
Bootstrapping for 1 proportion

BTTM
Bootstrapping for 2-sample means

RNMN
Randomization test for 1-sample mean

RNPR
Randomization test for 1-sample proportion

RNTM
Randomization test for 2-sample means

Design of experiments

BFFA
Analyze a binary response variable for a 2-level factorial design.

BGFA
Analyze a binary response variable for a general full factorial design.

BRSREG
Analyze a binary response variable for a response surface design.

BSCREEN
Analyze a binary response variable for a screening design.

MDESIGN
Modify a design that is in the worksheet.

VPREPROCESS
Pre-process responses for analyzing the variability of repeat or replicate measurements for a factorial design.

Obsolete commands

CONSTANT/NOCONSTANT
Use subcommands to for an individual analysis to control the estimation of the constant term.

DIR
List the names of files in a directory.

GPAUSE
Specify the number of graphs to display before you are prompted to save or discard open graphs. In Minitab 19.1, the number of graphs does not have a fixed limit.

GPRINT
Print a graph window. In Minitab 19.1, all output is in tabs instead of windows.

GVIEW
Open a .MGF format image. Minitab does not save images in this format anymore.

INSERT
Insert rows of data into the worksheet. Consider `WOPEN` and `READ`.

IW
Set the maximum width for input to the session window. Minitab 19.1 does not limit the input width in the Command pane.

JOURNAL/NOJOURNAL
Save session window lines to a text file. In Minitab 19.1, save the **History** pane.

MRESET
Restore environment settings to pre-macro conditions. In Minitab 19.1, restoration occurs at the end of every macro.

OW
Set the width of output in the session window. In Minitab 19.1, the output pane does not have a fixed width by number of characters.

PLUG/NOPLUG
Respond to errors from the macro processor. In Minitab 19.1, the macro processor stops when it encounters an error.

TITLE/NOTITLE
Display a title above session window output. In Minitab 19.1, use `MTITLE`/`ENDMTITLE` to add a title for an output tab and to group output on a single output tab.

Changes to subcommands

Opening and saving files

READ
The `FORMAT` subcommand does not use the T format.

RETRIEVE

The PORTABLE subcommand is obsolete. MTP format files are not compatible with Minitab 19.1.

The REPLACE/NOREPLACE subcommands are obsolete. If you save a file with the same filename, Minitab overwrites the file.

The GRAPH does not support the parameter MGF. Minitab no longer saves images in this format.

The PROJECT subcommand saves the project in the MPX format. The subsubcommand PASSWORD specifies a password that protects the file.

SAVE

The RELEASE subcommand accepts only 19 as a value. The tabbed output and other changes cannot be saved to an earlier version of Minitab.

The REPLACE/NOREPLACE subcommands are obsolete. If you save a file with the same filename, Minitab overwrites the file.

The GRAPH does not support the parameter MGF. Minitab does not save images in this format anymore.

The PROJECT subcommand saves the project in the MPX format. The subsubcommand PASSWORD specifies a password that protects the file.

Analysis of linear models

NLOG

OLOG

The TOLERANCE subcommand uses 1 parameter. Minitab uses the same value for all tolerances.

REGRESSION

The GPARETO subcommand makes a Pareto chart of the effects.

Changes to correlation

CORRELATION

The NOPVALUES subcommand is obsolete. P-values are in the new pairwise correlation table.

The following subcommands control the amount of output:

- TMETHOD
- TCORRELATION
- TPCORRELATION
- NODEFAULT
- GMPLOT

The following subcommands are for storage:

- SCORRELATION
- SCIS

The CONFIDENCE subcommand sets the confidence level for confidence intervals.

New subcommands

Creation of designed experiments

BBDESIGN

CCDESIGN

EVDESIGN

FDESIGN

FFDESIGN

MIXREG

OADESIGN

OPTDES

PBDESIGN

SCDESIGN

SLDESIGN

SPDESIGN
The DESIGN subcommand stores the design columns in the worksheet. The new subcommand replaces subcommands that required individual specification.

OPTDES
The COPY subcommand moves columns that are not part of the design, such as COVARIATES, to the new worksheet where you store rows for an optimal design.

ROBUST
The DOE subcommand stores the model information in a Minitab design object.

Analysis of linear models

FFAC

GFAC

GLM

GZLM

REGRESS

RSREGRESS

SCREEN

VFAC
The FINFORMATION subcommand specifies forward stepwise selection with information criterion.

Capability analysis

BWCAPA

CAPA

MCAPA

The UCPM and LCPM subcommands store one-sided confidence limits of the Cpm statistic.

The ONECI sub-subcommand to the CONFIDENCE subcommand specifies one-sided confidence intervals for the capability metrics.

The PPM subcommand replaces the percentage calculations with parts per million (PPM).

Changes to subcommands

FORMAT

The FORMAT subcommand, for commands such as READ, does not use the T format.

GSAVE

- REPLACE is obsolete.
- NOREPLACE is obsolete.

RELEASE

The RELEASE subcommand to SAVE accepts values of 19 or higher.

WOPEN

Changes to the subsubcommands of the subcommand FTYPE:

- XLSX is new.
- REPLACE is obsolete.
- NOREPLACE is obsolete.

WSAVE

Changes to the subsubcommands of the subcommand FTPYE:

- MINITAB accepts values of 19 or higher.
- XLSX is new.
- U8TEXT is new.
- U8CSV is new.
- REPLACE is obsolete.
- NOREPLACE is obsolete.
- WEBPAGE is obsolete.

Obsolete subcommands

FNUMERIC

The subsubcommand CULTURE to the subcommand CURRENCY is obsolete.

READ

The subcommand TAB is obsolete. Use WOPEN instead of READ to open tab-delimited data.

WSAVE

REPLACE and NOREPLACE are obsolete. If you save with the same filename, Minitab overwrites the file.

Commands that are unavailable in the web app

The following commands are not currently supported in exec or macro files in the Minitab web app:

Command	Description
CD	Display current directory or change current directory to another path
EXEC	Run a MTB file
EXIT	In macros, transfer control back to interactive Minitab; in execs, exit Minitab
GSAVE	Save a graph file
MYME	Invoke commands specified on customized My Menu
NEW	Create new project
NOOU	Stop writing output to an open/active text file
ODBC	Import data from a database file
OUTF	Save output to a text file
OW	Set character width of output
PAUS	Shift control to keyboard for interactive user input
PROF	Change profile
PYSC	Python script command
QUIT	Exit Minitab
READ; FILE	Import data from a text file into a worksheet (READ is allowed without FILE)
REST	Restart Minitab
RESU	Resume macro execution from interactive keyboard entry
RETR	Retrieve a project file in local macros
TSET; FILE	Import data from a text file (TSET is allowed without FILE)
SAVE	Save a project file in local macros
SET; FILE	Import data from an ASCII file (SET is allowed without FILE)
STOP	Exit Minitab
WOPE	Open worksheet
WSAV	Save worksheet
WRITE; FILE	Writes data in the specified columns or constants to a data file (WRITE is allowed without FILE)
XPPO	Send output to Microsoft® PowerPoint
XWOR	Send output to Microsoft® Word

Opening, Saving, and Printing Files

END: Session command for ending data input

END

Ends data input.

Type END following the last data line typed after READ (for information, go to READ data into a matrix on page 44 or READ data into columns on page 42), or SET on page 81. This ensures that any diagnostic messages concerning data lines will be printed before the next operation is carried out.

GSAVE: Session subcommand for saving a graph in a file

GSAVE ***"file_name"***

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB

TIF grayscale

TIF

TIF color

BMPB

BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

ODBC: Session command for importing data from a database file

ODBC
Imports data from a database file, such as one saved by Microsoft Access, Oracle, dBASE, Sybase, or SAS, into the Minitab worksheet.

With ODBC (open database connectivity), you can import a subset of data, such as data collected during a certain month, into the Minitab worksheet. ODBC adds data to the worksheet to the right of existing columns.

To use the ODBC session commands (for example, in a macro) use this method to identify the correct syntax:

1. Choose **File** > **Query Database (ODBC)** to query data.
2. When you have successfully retrieved the data you want, copy the corresponding command language from the **History** pane. (Open the **History** pane by pressing **Ctrl**+**K**.)
3. The COLUMNS subcommand is not created when you use the ODBC dialog boxes. If you are creating a local macro, add this subcommand yourself. Remember that in a local macro, columns must be declared before they are used as arguments in any commands. The ODBC session command can also query more than one table at a time.

CONNECT *"connection string"*
CONNECT specifies complete information pointing to the data source you want to query on your network. The text string argument can be very long.

SQLSTRING *"sqlstring"*
SQL specifies the query that selects a subset of the data. The text string can be up to 16384 characters in length.

COLUMNS k...k
COLUMNS specifies which columns of the Minitab worksheet should hold the data. k can be an integer from 1 to 4000.

The COLUMNS subcommand is required in local macros. In global macros or Execs, or when using Minitab interactively, executing the ODBC command without the COLUMNS subcommand places new data at the end of the global worksheet.

Note You can specify a range of columns with a colon, for example: `1:5`. Within Minitab or in a global macro you can also use stored constants that contain integers, for example `1:k1` where k1=5.

OUTFILE and NOOUTFILE: Session commands for saving a Minitab session in a text file

OUTFILE *"filename"*

OUTFILE K

OUTFILE saves your Minitab output in a text file. You can specify the filename as either the name of the file in double quotation marks, or as a stored text constant.

After OUTFILE is typed, Minitab sends a copy of the text from the output to a text file. OUTFILE is in effect until you type NOOUTFILE or you exit Minitab. If you type OUTFILE again, with the same file name, output is appended to the end of this file. To overwrite the file instead, use the REPLACE subcommand.

The file is a standard text file, which can be printed and edited by any editor or word processor. Unless you specify a different file extension, Minitab adds the extension LIS to the file name.

You can use OUTFILE to get a printout of your worksheet. Suppose your current worksheet contains data in the first 10 columns. For example, the following commands save a copy of your worksheet is to a file named Sales.LIS in your default directory.

```
OUTFILE "Sales".
PRINT C1-C10
NOOUTFILE.
```

File instructions

REPLACE

REPLACE causes the text to overwrite an existing file instead of appending to the file. Specify REPLACE on the line after OUTFILE.

```
OUTFILE "Sales";
REPLACE.
PRINT C1-C10
NOOUTFILE.
```

Closing the outfile

NOOUTFILE

NOOUTFILE closes an open outfile. If you type OUTFILE again with the same file name, input lines are appended to the same file unless you specify the REPLACE subcommand before NOOUTFILE.

PRINT: Session command for displaying columns, constants, or matrices in the output pane

PRINT *E...E*

Displays data in the output pane. You can display one or more (or any mixture of) columns, stored constants, or matrices.

If you mix columns, constants, and matrices in one PRINT command, they are listed in the following order: first all matrices and constants (in the order you specified), then all columns (in the order you specified). For example, the following command prints M1, K2, K1, and M2, then C1, C4, and C3.

```
PRINT C1 C4 M1 K2 C3 K1 M2
```

If a number is too large to fit in the space allowed (and FORMAT is not used), the number is displayed in exponential format.

FORMAT *(format statement)*

The FORMAT subcommand specifies where and how to print data on the output line. If you PRINT columns using FORMAT, only the data in columns is displayed, not column names. For more information, go to Valid format items on page 1269.

Minitab chooses the output format (i.e., number of decimal digits printed). If you want to control this yourself, use the FORMAT subcommand, or you can right-click on the output and select **Decimal Places**.

The following command prints numbers single-spaced, five numbers on each line, each number in a field of ten spaces.

```
PRINT C11-C15;
 FORMAT (5F10).
```

READ data into columns

READ C...C

Reads in data, row by row, that you type from the keyboard, or that you import from a text file. You cannot type comments on a data line when the FORMAT subcommand is used.

READ enters new data into columns, replacing any data already in those columns, if it exists. For information on entering data into a matrix, go to READ data into a matrix on page 44.

When you enter data manually, type `END.` after you enter your final value.

When you use READ, you can use a space or a comma to separate data entries. For example:

```
READ C1 C5.
1 2
3,4
END.
```

For details on using this command without subcommands to select data entry options, go to Using READ without subcommands on page 1266.

FILE *"filename"*

Inserts data from the specified text file. You may specify the filename as either the name of the file in double quotes, or a stored text constant. If the file has an extension other than DAT and/or if it is not in your current

directory, include the file extension and the path within the single quotation marks. For example, use the following command to read a copy of the file SALES.ASC stored in the subdirectory JANUARY underneath the directory SMITH on the C drive.

```
READ C1-C5;
 FILE "C:\SMITH\JANUARY\SALES.ASC".
```

FORMAT ***(format statement)***

Include a format statement, within parentheses to specify precisely how to enter data into the worksheet. The entire expression within the parentheses is repeated once for each record. For more information, go to Using READ with FORMAT on page 1266.

The FORMAT subcommand is useful when you want to skip over spaces, read data that have no spaces between them, insert decimal points in numbers, or read in text data, date/time data, or currency data.

Format items may be combined together. For example, the following command reads the name in the first 20 spaces of each data line into Name (C12), skips the next 10 spaces (spaces 21 through 30), then reads the number in space 31 into C1, the number in space 32 into C2, ..., the number in space 40 into C10.

```
NAME C12 'Name'
READ 'Name' C1-C10;
 FILE "MYDATA";
 FORMAT(A20, 10X, 10F1).
```

Minitab has a special date/time (DT) format which works as shown below. This says to read the date/time value in the first 8 spaces in the file into C1, and that the format of the date/time data in the file is m/d/yy.

```
READ C1;
 FILE "DATEDATA";
 FORMAT(DT8m/d/yy).
```

The following example shows the use of a decimal indicator, repeat factor in front of parentheses, and the slash.

```
READ C11-C15;
 FILE "EMPLOYEEDATA";
 FORMAT (F2.1, 2(1X,F3), F4/F2).
```

This example uses two data lines for every row read. From the first line, the value of C11 is in spaces 1 and 2. The first value is a whole number and the second value is in the tenths place. The format skips space 3. Then, C12 is read from spaces 4 to 6. The format skips space 7. Then, C13 is read from spaces 8 to 10, which repeats the pattern inside of the parentheses. C14 is read from spaces 11 to 14. In response to the /, reading moves to the second data line, and C15 is read from spaces 1 and 2. For more information, go to Valid format items on page 1269.

NOBS K

The NOBS subcommand specifies the number of observations (rows) to be inserted. If an END on page 39 subcommand or end-of-file is encountered before K observations are inserted, NOBS is ignored. NOBS is useful when you want to insert just the first portion of a file. It is also useful for Prompting a user for information on page 1260.

SKIP K

Tells Minitab to skip K lines at the top of the data file before beginning to add data into the file. This is most useful when you have one or more lines of text, such as column names and titles, at the top of a data file that you want to import into Minitab.

With READ; FILE only

DECIMAL works with READ when only when reading a file. DECIMAL does not work with READ if you type data.

DECIMAL ","

DECIMAL "."

Specifies a comma or period as a decimal separator.

READ data into a matrix

READ K K M

Puts numbers into a matrix. To input data to columns, go to READ data into columns on page 42. You can specify the filename as either the name of the file in double quotes, or a stored text constant. If the file has an extension other than DAT and/or if it is not in your current directory, include the file name extension and the path within quotation marks.

You can use either spaces or commas to separate the data in the matrix.

You must specify the dimension of the matrix in the READ command. The first K gives the number of rows, the second K the number of columns. The M is the matrix identifier for storage. If a file name is not used, READ is followed by data lines, each containing one row of the matrix. The following command creates the following matrix.

Command

```
READ  3  4  M2
1  2  3  4
5  6  7  8
9  10  11  12
END
```

Matrix

```
1  2  3  4
5  6  7  8
9  10  11  12
```

FILE *"filename"*

FILE K

Reads or inserts data from the specified text file.

RESTART: Session command for restarting

RESTART

Note You cannot use RESTART in a global macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Allows you to start over again.

RESTART erases the worksheet, cancels any controls in effect. RESTART also closes all open files such as OUTFILE.

RETRIEVE: Session command for retrieving a saved worksheet or project

RETRIEVE *"filename"*

RETRIEVE K

Note The menu command **File** > **Open** and the session command WOPEN on page 46 also open Minitab saved worksheets and Excel files (and many other types of files). They provide several useful options that are not available with RETRIEVE.

Use the main command by itself to retrieve a saved worksheet and add the file to the current project. With subcommands you can open a project or add one or more worksheets from a project to the current project. You can specify the filename as either the name of the file in double quotes or as a stored text constant.

If you omit the file name and the current folder contains a file named Minitab.MWX or Minitab.MTW, then Minitab opens that file.

Note You cannot use RETRIEVE in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

PROJECT

Note You cannot use PROJECT in a global macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Specifies that the file after RETRIEVE is a Minitab project file (MPX, MPJ). If you do not want the prompt, use the SAVE command with the PROJECT subcommand before you use the RETRIEVE command.

PASS *"password"*

To retrieve a password-protected file, specify the password.

SAVE: Session command for saving a worksheet or project

SAVE [K]

Note You cannot use SAVE in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Saves a worksheet or project. You can specify the filename as either the name of the file in double quotation marks, or a stored text constant. When no subcommands are specified, worksheet (MWX) is the default file type for SAVE.

A saved worksheet file contains all data, stored constants, matrices, column names, and missing value information. Minitab automatically replaces an existing file if you save with the same filename.

You can open saved worksheets with RETRIEVE on page 45 or WOPEN on page 46.

PROJECT

Specifies to save as a Minitab Project file (MPX).

WSONLY
Use the subcommand WSONLY to save only worksheets with the saved project.

PASSWORD "password"
Specify a password to use to open the file. Enter the password within double quotation marks. To remove password protection, use the argument "".

RELEASE K
Specifies the earliest version of Minitab that can open the file. For Minitab 19 and higher, 19 is the earliest valid argument. You cannot specify a version newer than the version of Minitab that you have.

STOP: Session command for closing Minitab

STOP
Closes Minitab.

WOPEN: Session command for opening a worksheet

WOPEN K

Note You cannot use WOPEN in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Opens a worksheet. For example, to open a data set named PULSE, enter the following command.

```
WOPEN "PULSE"
```

When you open a file, you copy the contents of the file into the current Minitab project. Any changes that you make to the worksheet while in the project do not affect the original file.

FTYPE
Specifies the type of the file to open. Use one of the following subcommands for FTYPE:

MINITAB

XLSX (For Excel files with a .xlsx extension)

EXCEL (For Excel files with a .xls extension)

XMLEXCEL

TEXT

CSV

You must specify the file type, such as EXCEL, to use some of the other subcommands, such as VNAMES. If you do not specify the file type, then Minitab considers the file extension. For example, if you enter `WOPEN "Mywork.xls"` but do not specify a file type, then Minitab opens the file as an Excel file. However, if that Excel file were saved with the name Mywork.abc, Minitab would not recognize the file as an Excel file, and would display an error message or open the file incorrectly.

When you open a Minitab worksheet (with `FTYPE; MINITAB`), you do not need to specify a release number.

FIELD

FIELD allows you to specify how fields are delimited.

Use one of the following subcommands to specify what character denotes the sections of data to be placed into each field. If FIELD is not specified, Minitab uses the default the default delimiter for the file type.

TAB

COMMA

SEMICOLON

SPACE

PERIOD

CUSTOM "K"

TDELIMITER

Specifies how text fields are delimited.

DOUBLEQUOTE

Use when the columns of data are separated by a double quotation mark.

SINGLEQUOTE

Use when the columns of data are separated by a single quotation mark.

NONE

Use to have text that is separated by a blank entered into its own column.

CUSTOM K

Use if the columns are separated by a character other than those listed above and enter the character as K.

DECSEP

Specifies how decimals are separated.

COMMA

Specifies that decimals are separated by a comma.

PERIOD

Specifies that decimals are separated by a period.

MISSING

Specifies how to write out a missing value.

TEXT K K

Denotes text columns, with the first K as the "missing" marker in the file's original form, and the second K as the one or more words, characters, or spaces that you want Minitab to replace these markers with. The default Minitab marker for missing text is a space.

NUMERIC 'K' K

Denotes numeric columns, with the first K as the missing marker in the file's original form, and the second K as the digits, characters, or spaces that you want Minitab to replace these markers with. The default Minitab marker for missing numeric data is an asterisk (*). The missing value symbol (first K) must be enclosed in single quotation marks.

DATA

Specifies where to begin reading data in the file to be opened.

IGNOREBLANKROWS

Skips over blank data rows. If you don't use this command, Minitab reads blank data rows as missing values.

EQUALCOLUMNS

Adds missing values (*) to shorter columns to so that all columns have the same number of rows.

CLEAN

Removes nonprintable characters and extra spaces from text columns.

CASE

Corrects case mismatches in text columns by applying the capitalization of the first occurrence of a text value to all matching values in the column.

SHEET K

Controls the way the K^{th} worksheet in the file is read. SHEET can be used one time for each worksheet in the file.

EXCLUDE

Stops the worksheet from being imported.

VNAMES K

Specifies the row number K that contains the variable names. −1 indicates no variable names. 0 indicates that the first row that is being imported is the variable names row.

FIRST K

Specifies the row number K to begin reading the data. If you do not choose a variable names row, Minitab begins reading data from the first available row in the file.

NROWS K

Specifies the number of data rows. Enter a number to read a portion of the file. For example, if you enter 50, Minitab reads only the first 50 rows beginning with the first row of data. By default, Minitab reads all rows from the file into the current worksheet.

VARIABLE K

Controls the way the K^{th} variable in the worksheet is read. It can be used once for each variable in the worksheet.

EXCLUDE

Do not import this variable.

NAME K

Assigns a name to the variable. Names can be up to 31 characters long, and can include any letters and numbers except the symbol ' (single quotation mark) or # (pound sign). Names cannot begin or end with a blank or consist entirely of the symbol * (asterisk). You also can name your variables directly after the worksheet is open.

TEXT

Specifies that the data type is text.

NUMERIC

Specifies that the data type is numeric.

DATETIME
Specifies that the data type is date/time.

COLUMN C
Specifies which column of the worksheet (C) the variable is placed in.

WORKSHEET: Session command for making a worksheet active, for closing a worksheet, or for renaming a worksheet

WORKSHEET K

You cannot use WORKSHEET in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Specifies K as the active worksheet. If no worksheet is specified, then session commands work on the active worksheet.

When you are working in Minitab, any session command you use works on the active worksheet. The active worksheet is the worksheet associated with the active Data tab.

You can also click a worksheet to activate it. If no worksheet is active, the session command acts on the worksheet that was most recently active.

Note A worksheet can contain up to 4000 columns, 1000 constants, and up to 10,000,000 rows depending on how much memory your computer has.

If you run subcommands, only the last subcommand works. To rename and close a worksheet, you must use more than one instance of the main session command.

CLOSE
Closes the worksheet.

NOPROMPT
Specifies not to prompt to save when saving the worksheet.

CURRENT
Specifies to make the named worksheet active.

RENAME K

RENAME *"text"*
Renames the worksheet.

WRITE: Session command for writing data to the screen or a data file

WRITE ***E...E***

Writes data in the specified columns or constants to the screen or to a data file (also called text or ASCII file). You can specify the filename as either the name of the file in double quotes, or a stored text constant.

WRITE exports your data to a data file which you can import into other applications, print on your printer, or enter into Minitab with READ (for more information, go to READ data into columns on page 42 or READ data into a matrix on page 44) or SET on page 81. WRITE prints the columns vertically, very close together, with no column names or row numbers and prints constants horizontally, whether they are stored numeric or text constants.

If you omit the file name, WRITE displays the columns or constants on your screen. If you specify a file name but omit the extension, WRITE adds the default extension DAT.

Use the FORMAT subcommand to write data with a fixed format, for example, columns not separated by spaces.

By default, if you WRITE to a file that already exists, Minitab overwrites the file. You can use the subcommand NOREPLACE to prevent the overwriting of the original file.

WRITE normally creates a data file. It can also be used to print columns or constants on your screen or on paper. The output is very compact. With columns, there is no header giving the column name, and there are no row numbers on the left. Columns are always output vertically. Constants are output horizontally, whether they are stored numeric or text constants.

A data file created with the WRITE command can be transferred to different computer types and read by other programs. On most computers, the default file name extension is DAT.

- The format of the output is adjusted to make the data as compact as possible. Thus, the number of columns that can be put on one line varies with the data.
- If a format is not specified, and if very wide text items are to be printed (e.g., 80 wide), and if the output width is not wide enough to print the entire item, the items will be truncated on the right.
- When writing columns of unequal length to a file, Minitab makes them equal by adding missing value symbols (*) to the short columns.

Note The menu command **File** > **Save Worksheet As** and the session command WSAVE on page 52 can also save the current Minitab worksheet as a text file, and many other file types as well. It also provides several useful options not available with WRITE.

FILE ***"filename"***

Specify to write to a data file (also called text or ASCII file). You may specify the filename as either the name of the file in double quotes, or a stored text constant. If you specify a file name but omit the extension, WRITE adds the default extension DAT. Use the subcommands UTEXT and TEXT to specify Unicode or ANSI text for the file.

UTEXT

Saves the file in Unicode format.

TEXT

Saves the file in text format.

FORMAT

A format specification, similar to a Fortran format, may be given on this subcommand. For more information, go to Valid format items on page 1269.

The format specifies where the data will appear on the output lines, and the number of decimal places printed. FORMAT works the same way with WRITE as it does with READ (with the exception of the currency format, which works with READ but not WRITE). For more information, go to READ data into columns on page 42 or READ data into a matrix on page 44.

The following table shows how 123.456 is output with different formats.

Format	Output
F7.3	123.456
F7.2	123.46
I7	123
E11.6	1.23456E+02
G7	123.456

If you create a file with WRITE, using a FORMAT, then the READ command (for more information, go to READ data into columns on page 42 or READ data into a matrix on page 44) can read the file with the same format.

To create a file that READ can read without a format (or other programs can read without a format), make sure that there is at least one space between numbers. One way to do this is with the X format item. For example, the following command creates a file which can be read without a format.

```
WRITE C1-C5;
    FILE "FileA";
 FORMAT ( 5(1x, F11) ).
```

Missing numeric data are output as * in the right-most space of the field. Missing text data are output as blanks.

You must WRITE columns of equal length.

You can write columns that contain text data with or without using a FORMAT subcommand. You must, however, use a FORMAT subcommand to READ the file.

Column format

TAB

TAB saves the specified columns, separated by tabs, in a data (also called text or ASCII) file. Do not use FORMAT with tab-delimited data.

NONAMES

If you do not want to include the column names in the file, use the NONAMES subcommand in addition to TAB.

File instructions

NOREPLACE

NOREPLACE prevents you from overwriting a file and is most useful in macros. When you specify NOREPLACE and you have an existing file with the same name, Minitab generates an error message and aborts the command or exits the macro.

WSAVE: Session command for saving a worksheet file

WSAVE ***"filename"***

Saves all the worksheet data in a file. Use this command to rename your worksheet or save it to a new location.

If your current worksheet is untitled ("Worksheet1" appears in the tab of the worksheet), you can use WSAVE to specify a name for your saved worksheet file.

FTYPE

Specifies a file format. The sub-subcommands (MINITAB, XLSX, and so on) are the file formats.

You must specify an FTYPE; FORMATNAME combination to use some of the other subcommands, such as VNAMES.

If you do not use FTYPE when saving a file, Minitab saves the file as a Minitab worksheet for the version that you have.

MINITAB K

Specifies the Minitab worksheet file type.

When you use the MINITAB file type, specify the earliest version of Minitab that can open the worksheet. For Minitab 19 and higher, 19 is the earliest valid argument. You cannot specify a version newer than the version of Minitab that you have.

XLSX

Specifies the Excel file type with a .xlsx extension.

EXCEL

Specifies the Excel file type with a .xls extension.

XMLEXCEL

Specifies the XML Excel file type.

TEXT

Specifies the text file type.

UTEXT

Specifies the unicode text file type.

CSV

Specifies the comma-separated file type.

UCSV

Specifies the unicode comma-separated file type.

NOVNAMES

Saves a file from Minitab into another format without saving the variable names within the file.

DECSP

Specifies the form of decimal separation when you save a text file. The sub-subcommands give the decimal separator.

COMMA
Specifies a comma decimal separator.

PERIOD
Specifies a period decimal separator.

FIELD
Specifies how fields are delimited. Choosing one of the options chooses what character denotes the sections of data to be put into each field. If FIELD is not specified, Minitab uses the default delimiter for the file type. The sub-subcommands give the delimiter.

TAB
Specifies the tab as the delimiter.

COMMA
Specifies the comma as the delimiter.

SEMICOLON
Specifies the semicolon as the delimiter.

PERIOD
Specifies the period as the delimiter.

SPACE
Specifies the space as the delimiter.

CUSTOM K
Specifies a custom delimiter.

TDELIMITER
Specifies how text fields are delimited.

DOUBLEQUOTE
Use when the columns of data are separated by a double quotation mark.

SINGLEQUOTE
Use when the columns of data are separated by a single quotation mark.

NONE
Use to have text that is separated by a blank entered into its own column.

CUSTOM K
Use if the columns are separated by a character other than those listed above and enter that character as K.

MISSING
Specifies how to write out a missing value. The sub-subcommands give the values.

TEXT K K
TEXT denotes text columns, with the first K as the "missing" value symbol in the file's original form, and the second K as the word(s), character(s), or space(s) you want Minitab to replace these symbols with. The default Minitab symbol for missing text is a space.

NUMERIC K K

NUMERIC denotes numeric columns, with the first K as the "missing" value symbol in the file's original form. The default Minitab symbol for missing numeric data is an asterisk (*). The missing value symbol (first K) must be enclosed in single quotation marks, '*'. The second K can be a number or '*' and is how the symbol appears in the output file. In Excel files, the missing value symbol becomes an empty cell.

XPPOINT: Session command for sending output to Microsoft PowerPoint

XPPOINT

Sends Minitab output and graphs to Microsoft PowerPoint so that you can create reports and presentations.

All of Minitab's graphs and output can be sent to a PowerPoint presentation. In PowerPoint, text and titles appear in the font that is used in Minitab.

Note When you send a graph to PowerPoint, Minitab's Embedded Graph Editor is not available.

APPEND

Exports output into an open Microsoft PowerPoint presentation. The output is starts on a new slide after the currently active slide.

XWORD: Session command for sending output to Microsoft Word

XWORD

Sends Minitab output and graphs to Microsoft Word so that you can create reports and presentations.

All of Minitab's graphs and output can be sent to a Word document. In Word, text and titles appear in the font that is used in Minitab.

Note When you send a graph to Word, Minitab's Embedded Graph Editor is not available.

APPEND

Exports output into an open Microsoft Word document. The output is placed after the cursor location in the Word document.

Other Session Commands

PYSC: Session command for running a Python script

PYSC ***["filename.py"] ["Args"...]***

Runs the `Python` script that you specify.

The default file extension for `Python` scripts is .PY. If the file extension is .PY, you do not need to type the file extension.

The optional argument `Args` lets you pass arguments to the `Python` script through `sys.argv[1:]`. `Args` can be any text values separated by a space. Enclose arguments in quotation marks. The default value is `None`, which means that the script does not receive any arguments.

For more information on integrating `Python` and Minitab, go to the Minitab Support Center.

NOSERR

Specifies not to display text from the `stderr` file in the Output pane in Minitab. For example:

```
PYSC "test.py";
NOSERR.
```

SOUT

Specifies to display text from the `stdout` file in the Output pane in Minitab. For example:

```
PYSC "test.py";
SOUT.
```

RSCR: Session command for running a R script

RSCR ***["filename.R"] ["Args"...]***

Runs the `R` script that you specify.

The default file extension for `R` scripts is .R. If the file extension is .R, you do not need to type the file extension.

The optional argument `Args` lets you pass arguments to the `R` script through `commandArgs(trailingOnly = TRUE)`. `Args` can be any text values separated by a space. Enclose arguments in quotation marks. The default value is `Null`, which means that the script does not receive any arguments.

For more information on integrating `R` and Minitab, go to the Minitab Support Center.

NOSERR

Specifies to not display text from the standard error (`message()`, `warning()`, or `stop()`) console output in the Output pane in Minitab. The warning console output is where `R` error messages display when you run your code in an `R` integrated development environment. For example:

```
RSCR "test.R";
NOSERR.
```

SOUT

Specifies to display text from the standard console output in the Output pane in Minitab. The standard console output is where the results of commands like `print()` display in an `R` integrated development environment. For example:

```
RSCR "test.R";
SOUT.
```

NAME: Session command for assigning names to columns, stored constants, and matrices

NAME ***E "name"...E "name"***

Assigns names to columns, stored constants, and matrices.

Names must not do the following:

- Contain more than 31 characters
- Begin or end with a blank
- Include the symbols ' or #
- Start with or consist entirely of the symbol *
- Be repeated. You cannot use the same name for two variables (columns, stored constants or matrices) in the same worksheet.

You can refer to a named column, constant, or matrix by either number or name, for example, M1 or "Inverse". Any time you use a name, enclose the name in double quotation marks. Minitab always prints variable names, if they exist, on all output involving that variable. You can change the name of a variable by using another NAME command.

The following example names C1, K1, and K2.

```
NAME C1"Height" K1"Min" K2"Max"
```

The following example erases the name of C1.

```
NAME C1  "  "
```

You can use both upper and lower case letters in a name. When you use the name in a Minitab command, Minitab considers upper and lower case letters equivalent. When a name is printed in the output, however, the upper and lower case letters are used. For example, the following commands name the C1 column Height and print the data.

```
NAME C1"HEIGHT"
 PRINT "HEIGHT"
```

Note You also can type column names into the worksheet as an alternative to the NAME command.

ABORT: Subcommand for exiting a multi-line command

ABORT

Use ABORT as the next subcommand to exit from a multi-line command without executing it.

HELP: Session command for opening this guide

HELP

Opens the Session Commands guide. The Session Commands guide provides summaries of commands and specific information on each command and subcommand. You can use session commands as an alternative to using menu commands, or as a way to build macros for repetitive functions.

Dynamic Data Exchange

XDACTIVATE: Session command for activating a link

XDACTIVATE ***"application" "topic" "item"***

Activates a previously deactivated client link. When you deactivate a link using XDDEACTIVATE on page 60, Minitab ignores/discards any queued transactions or incoming transactions until you reactivate the link.

You can use XDACTIVATE in execs and global macros, but not in local macros.

You also can pause activity globally.

application

Application is the name of a program that can participate in a DDE transaction. Usually, application is the name of an EXE file that starts the program, without the .exe extension. For example, in the case of Minitab, Mtb.exe starts the program, so for the application argument use `Mtb`.

topic

Topic is a name that depends on the type of application. In applications that use documents or data files, the topic is often the name of the file. For example, the topic for a Microsoft Word document might be mywork.doc. The file name can also include a path, as in c:\mywork\mywork.doc.

If one file can contain many documents or subwindows, the Topic can be the name of the file in brackets, followed by the name of the document or window. For example, Minitab projects can contain many worksheet, so a typical Topic would be [Minitab] worksheet 1.

Other applications have their own names for topics.

Tip It is a good idea to save files from other applications before establishing any links to or from them. If you establish a link before saving the file, you may need to change the link name from something like "untitled" to a file name. Likewise, before linking from Minitab, rename a new worksheet from the default of "Worksheet 1" to something meaningful, such as "First Quarter."

item

Item is a name that depends on the type of application. In Minitab, the item always specifies a row/column location or rectangular area in the form R4C1:R4C2. You can use all of Minitab column 1 as an item by specifying C1. A single cell (row 4, column 1) is R4C1. Most spreadsheets use close variants of the R4C1:R3C2 format for items.

XDADD: Session command for adding a new link

XDADD ***"application" "topic" "item"***

Establishes a new DDE link between the specified application program and Minitab. With this command, Minitab acts as a client, receiving data from the other application.

You often specify XDADD with either the APPEND or REPLACE subcommands to tell where to put the data in the Minitab worksheet, as shown below.

```
XDADD "Minitab" "[myproj.mpj]pulse" "C1";
  REPLACE C2;
  PERFORM 2;
  COMMAND "LET C2=4*C1".
```

When another program supports **Edit** > **Paste Link**, you can find out the application, topic, and item names by copying a range from the other application and pasting it into the worksheet in Minitab. The type of link established by XDADD is often referred to as a hot link.

You can use XDADD in global macros and exec macros, but not in local macros.

Data

DATA

Specifies whether to add (append) the data to the end of the columns that you specify or to replace the entire contents of the columns in the worksheet with new data from the item you specify in the XDADD or XDGET command. The column argument specifies the starting, or anchor column for the data in the item. When the item has more than one column, other columns are also included.

APPEND [C]

Specifies to add the data to the end of the columns that you specify.

REPLACE [C] (default)

Specifies to replace the entire contents of the columns in the worksheet with new data from the item you specify in the XDADD or XDGET command.

Link status

ACTIVE

Specifies to open the new link in the active state. An active link starts receiving data from the other application immediately.

INACTIVE

Specifies to open the new link in the inactive state. An inactive link does not receive data until you activate it.

Command arguments

PERFORM K

Specifies the action to take on each occurrence of a data transfer from the server application. The argument values and associated actions are shown in the following table.

Value of K	Description
0	Update data and execute commands.
1	Update data only.
2	Execute commands only.

COMMAND *"command"*

Specifies one Minitab command to execute upon each occurrence of a data transfer from the server application. The command string can contain up to 128 characters (without line breaks) and include one Minitab command or %macro. Do not use a period at the end.

Use a %macro or the dialog box if you need to use multiple commands.

If the command includes single quotation marks ('), enclose the single quotation marks with double quotation marks ("). You can only use a single COMMAND subcommand for each link.

This subcommand works differently than the **Commands** field in the **Add New Links** dialog box.

Link priority

PRIORITY K

Specifies the priority of the link relative to other links. The priority can be any integer from 1 (the highest priority) to 32 (the lowest priority). Priority matters when more than one link attempts a transfer at the same time, or when several queued transactions are waiting for processing. Minitab always completes transfers on links that have higher priority first, then processes data from links with lower priority. Links with the same priority are transferred on a first-in first-out basis. The default priority is 16.

XDDEACTIVATE: Session command for deactivating a client link

XDDEACTIVATE ***"application" "topic" "item"***

Allows you to pause (deactivate) activity on a client link. When you deactivate a link, Minitab ignores/discards any queued transactions or incoming transactions until you use XDACTIVATE on page 58 to reactivate the link.

You can use XDDEACTIVATE in execs and global macros, but not in local macros.

You can also pause activity globally.

application

Application is the name of a program that can participate in a DDE transaction. Usually, application is the name of an EXE file that starts the program, without the .exe extension. For example, in the case of Minitab, Mtb.exe starts the program, so for the application argument use `Mtb`.

topic

Topic is a name that depends on the type of application. In applications that use documents or data files, the topic is often the name of the file. For example, the topic for a Microsoft Word document might be mywork.doc. The file name can also include a path, as in c:\mywork\mywork.doc.

If one file can contain many documents or subwindows, the Topic can be the name of the file in brackets, followed by the name of the document or window. For example, Minitab projects can contain many worksheets, so a typical Topic would be [Minitab] worksheet 1.

Other applications have their own names for topics.

Tip It is a good idea to save files from other applications before establishing any links to or from them. If you establish a link before saving the file, you may need to change the link name from something like "untitled" to a file name. Likewise, before linking from Minitab, rename a new worksheet from the default of "Worksheet 1" to something meaningful, such as "First Quarter."

item

Item is a name that depends on the type of application. In Minitab, the item always specifies a row/column location or rectangular area in the form R4C1:R4C2. You can use all of Minitab column 1 as an item by specifying C1. A single cell (row 4, column 1) is R4C1. Most spreadsheets use close variants of the R4C1:R3C2 format for items.

XDEXEC: Session command for executing a command in a remote application

XDEXEC ***"application" "command"***

Allows you to execute a command in a remote application.

You can use XDEXEC in any type of macro.

application

The application argument is the DDE application name that is defined by the other application.

Application is the name of a program that can participate in a DDE transaction. Usually, application is the name of an EXE file that starts the program, without the .exe extension.

command

If the other application allows more than one command per line, you can use more than one command in the command string. If the server does not allow more than one command, consider writing a macro in the server language, or using several XDEXEC commands in a row to achieve the same result. Use double quotation marks inside the command string if single quotation marks are required in the remote application.

Note If the command that you want to execute already contains double quotation marks, then put the command and application in single quotation marks. For example, because the syntax for executing an Excel macro requires double quotation marks, the correct syntax is `XDEXEC 'EXCEL' '[Run("macro")]'`.

XDGET: Session command for performing a one-time data transfer

XDGET ***"application" "topic" "item"***

Allows you to do a one time data transfer (often referred to as a cold link). You can use XDGET in place of copying and pasting or in any type of macro.

application

Application is the name of a program that can participate in a DDE transaction. Usually, application is the name of an EXE file that starts the program, without the .exe extension. For example, in the case of Minitab, Mtb.exe starts the program, so for the application argument use `Mtb`.

topic

Topic is a name that depends on the type of application. In applications that use documents or data files, the topic is often the name of the file. For example, the topic for a Microsoft Word document might be mywork.doc. The file name can also include a path, as in c:\mywork\mywork.doc.

If one file can contain many documents or subwindows, the Topic can be the name of the file in brackets, followed by the name of the document or window. For example, Minitab projects can contain many worksheets, so a typical Topic would be [Minitab] worksheet 1.

Other applications have their own names for topics.

Tip It is a good idea to save files from other applications before establishing any links to or from them. If you establish a link before saving the file, you may need to change the link name from something like "untitled" to a file name. Likewise, before linking from Minitab, rename a new worksheet from the default of "Worksheet 1" to something meaningful, such as "First Quarter."

item

Item is a name that depends on the type of application. In Minitab, the item always specifies a row/column location or rectangular area in the form R4C1:R4C2. You can use all of Minitab column 1 as an item by specifying C1. A single cell (row 4, column 1) is R4C1. Most spreadsheets use close variants of the R4C1:R3C2 format for items.

APPEND and REPLACE

Specifies whether to add (append) the data to the end of the columns that you specify or to replace the entire contents of the columns in the worksheet with new data from the item you specify in the XDADD or XDGET command. The column argument specifies the starting, or anchor column for the data in the item. When the item has more than one column, other columns are also included.

APPEND [C]

Specifies to add the data to the end of the columns that you specify.

REPLACE [C] (default)

Specifies to replace the entire contents of the columns in the worksheet with new data from the item you specify in the XDADD or XDGET command.

XDREMOVE: Session command for deleting an established link

XDREMOVE *"application" "topic" "item"*

Deletes an established client link.

This command deletes the specified client link immediately when Minitab accepts the command. When a link is deleted, all existing queued requests are processed (that is transferred into the worksheet), but no new ones are added.

You can use XDREMOVE in execs and global macros, but not in local macros.

application

Application is the name of a program that can participate in a DDE transaction. Usually, application is the name of an EXE file that starts the program, without the .exe extension. For example, in the case of Minitab, Mtb.exe starts the program, so for the application argument use `Mtb`.

topic

Topic is a name that depends on the type of application. In applications that use documents or data files, the topic is often the name of the file. For example, the topic for a Microsoft Word document might be mywork.doc. The file name can also include a path, as in c:\mywork\mywork.doc.

If one file can contain many documents or subwindows, the Topic can be the name of the file in brackets, followed by the name of the document or window. For example, Minitab projects can contain many worksheets, so a typical Topic would be [Minitab] worksheet 1.

Other applications have their own names for topics.

Tip It is a good idea to save files from other applications before establishing any links to or from them. If you establish a link before saving the file, you may need to change the link name from something like "untitled" to a file name. Likewise, before linking from Minitab, rename a new worksheet from the default of "Worksheet 1" to something meaningful, such as "First Quarter."

item

Item is a name that depends on the type of application. In Minitab, the item always specifies a row/column location or rectangular area in the form R4C1:R4C2. You can use all of Minitab column 1 as an item by specifying C1. A single cell (row 4, column 1) is R4C1. Most spreadsheets use close variants of the R4C1:R3C2 format for items.

Manipulating and Calculating Data

Calculator

Calculations

ADD: Session command for addition

ADD *E...E E*

Adds E and E and stores in E. E can be any column, any constant, or matrix. For columns and constants, ADD allows up to 50 arguments. If any element of a row is missing, the result is set to missing *. If the operation is impossible, the result is also set to missing.

EIGEN: Session command for calculating eigenvalues

EIGEN *M* C *[M]*

Calculates eigenvalues (also called characteristic values or latent roots) and eigenvectors for a symmetric matrix.

The eigenvalues are stored in decreasing order of magnitude down the column. The eigenvectors are stored as columns of the matrix. The first column corresponds to the first eigenvalue (largest magnitude), the second column to the second eigenvalue, and so on.

LAG: Session command for calculating the lags of a column

LAG [K] C C

Calculates the lags of a column and stores them in a new column.

Moves the row elements of a column down K rows, where K is the lag specified, storing the result in a new column of the same length. There will be K missing value symbols, *, at the top of the output column. The output column has the same number of rows as the input column, so the last K values from the input column are not lagged. If K is omitted, then K = 1 is used.

If C1 contains Z1, Z2, ..., Zn, then `LAG K C1 C2` puts asterisks (*) into rows 1 through K of C2, and puts $Z_i - Z_{i-k}$ into row i, K + 1 < i < n.

LET: Session command for correcting a number in a worksheet or performing arithmetic

LET C [K] = K

Use to correct a number in the worksheet. You can change a single number with the LET command. For example, `LET C1[5] = 28.6` puts 28.6 into row 5 of C1. The rest of C1 remains the same.

LET = *expression*

Use to perform arithmetic with an algebraic expression. The expression can contain arithmetic operations, comparison operations, logical operations, and functions. The functions that you can use with LET are described below.

Note In an expression, you cannot use a hyphen to specify a range of columns or constants. For example, Minitab would interpret C1-C4 as C1 minus C4 and 1-10 as 1 minus 10.

You can also access an individual row in a column. For arguments, you can use columns, stored constants, or numbers. You cannot use matrices. You cannot use extra text on any session command, including LET, except after the # symbol.

For arithmetic operations, if any element of a row is missing, the result is set to missing. If an operation is impossible, such as division by zero, the result is set to missing. You can have up to nine nested parentheses.

Examples of LET

```
LET C1 = (C2 + C3)*10 - 60

LET C1 = C1 - MEAN(C1)

LET K1 = 5.3

LET K2 = MEAN(C10)/STDEV(C1)

LET C5 = (C1 < 5)

LET K2 = C1[28]

LET C1[28] = 2.35

LET C5[15] = "blue"
```

Arithmetic functions to use with LET

CEIL *(E, K)*

Rounds numbers up. In the first argument, specify the number you want rounded. In the second argument, specify the number of decimals to round to. If K = 0 (the default), the number is rounded to the nearest integer greater than or equal to the number. If K > 0, the number is rounded up to the specified number of decimal places after the decimal point. If K < 0, the number is rounded up to 1- the specified number of place values to the left of the decimal point.

For example:

CEILING (2.136, 0) equals 3

CEILING (2.136, 1) equals 2.2

CEILING (2.136, 2) equals 2.14

CEILING (–2.136, 1) equals –2.1

CEILING (253.6, –1) equals 260

CEILING (253.6, –2) equals 300

COMBINATIONS *(E, E)*

Calculates the number of combinations of n items chosen k at a time. In the first argument, specify the number of items. In the second argument, specify the number to choose. The number of items must be greater than or equal to 1, and the number to choose must be greater than or equal to 0. Arguments can be columns or constants. Missing values are not allowed.

FACTORIAL *(E)*

Calculates the factorial of a number, the product of all the consecutive integers from 1 to the number, inclusive. The value of the number must be greater than or equal to 0. Missing values are not allowed.

FLOOR *(E, K)*

Rounds numbers down. In the first argument, specify the number you want rounded. In the second argument, specify the number of decimals to round to. If K = 0 (the default), the number is rounded to the nearest integer less than or equal to the number. If K > 0, the number is rounded down to the specified number of decimal places after the decimal point. If K < 0, the number is rounded down to 1- the specified number of place values to the left of the decimal point.

For example:

FLOOR (2.136, 0) equals 2

FLOOR (2.136, 1) equals 2.1

FLOOR (2.136, 2) equals 2.13

FLOOR (–2.136, 1) equals –2.2

FLOOR (253.6, –1) equals 250

FLOOR (253.6, –2) equals 200

FTC *(E)*

Performs the Freeman Tukey transformation to stabilize variance for Poisson data. Requires one argument, which must be a column or stored constant that contains nonnegative integers.

FTP *(E, E)*

Performs the Freeman Tukey transformation to stabilize variance for binomial data. Requires one argument for number of trials and one for number of successes. Each argument can be a column or stored constant. Trials must be a positive integer, and successes must be an integer between 0 and n inclusive.

GAMMA *(E)*

Calculates the gamma function where E is the specified shape parameter (the number you want to take the function of).

IGAMMA *(E, E)*

Calculates the incomplete gamma function. In the first argument, specify the upper limit of the integral. In the second argument, specify the shape parameter (the number you want to take the function of).

LNGAMMA *(E)*

Calculates the natural log of the gamma function. E is the shape parameter.

MOD *(E, E)*

Stores the remainder after a number is divided by a divisor. In the first argument, specify the number. In the second argument, specify the divisor. Minitab calculates the value using the formula: m – (n * FLOOR (m/n)) where m is a number and n is the divisor.

PERMUTATIONS *(E, E)*

Calculates the number of permutations of n things taken k at a time. In the first argument, specify the number of items. In the second argument, specify the number to choose. The number of items must be greater than or equal to 1, and the number to choose must be greater than or equal to 0. Missing values are not allowed.

Column functions to use with LET

DIFFERENCES *(C, K)*

Calculates row-by-row differences between the numeric values in a column. In the first argument, specify the column. Minitab subtracts from each row the element K rows above, where K is the lag you specify, and stores the differences in a new column. If you don't specify a value for lag, the differences are computed between consecutive rows (lag = 1). The first K rows of the new column will contain the missing value symbol, *.

LAG *(C, K)*

Moves the row elements of a column down K rows, storing the result in a new column of the same length. If no value is specified, the default lag (K =1) is used. There will be K missing value symbols, *, at the top of the output column. The output column has the same number of rows as the input column, so the last value from the input column is not lagged.

NSCORES *(C)*

Calculates and stores normal scores. This command is used mainly to produce normal probability plots and perform various tests. Here is an example of normal scores.

```
LET C2 = NSCORES(C1)
```

Loosely speaking, -1.18 is the smallest value you would expect to get if you took a sample of size 5 from a standard normal (i.e., the expected value of the first order statistic), and -0.50 is the second smallest value you would expect to get (i.e., the expected value of the second order statistic), and so on.

C1	C2
1.1	0.0
0.1	-1.18
2.3	1.18
1.8	0.50
0.9	-0.50

Minitab does not calculate the expected values of the order statistics exactly, but uses percentage points as an approximation. Thus, if there are n data values, Minitab puts f - *1*[(i - 3 / 8) / (n + 1 / 4)] next to the i*th* smallest data value, where f -*1*(x) is the inverse cumulative distribution function of the standard normal.

If several observations are equal, they are all given the same normal score. This is calculated using the average of their ranks. The command NORMPLOT provides additional information. The output from NORMTEST includes a normal probability plot, relevant statistics about the data, and test statistics useful for testing a hypothesis of normality.

RANK (C)

Calculates and stores the ranks of the input column. Assigns the numeral 1 to the smallest value, the numeral 2 to the second smallest value, the numeral 3 to the third smallest value, and so on. Ties are assigned the average rank.

Note RANK works only with numeric columns.

The following command language ranks the values in C1 and puts the ranked values in C2.

```
RANK C1 C2
```

Before ranking

C1
0.5
1.0
1.5
1.0
2.0
0.0

After ranking

C1	C2
0.5	2.0
1.0	3.5
1.5	5.0
1.0	3.5
2.0	6.0
0.0	1.0

SORT (C)

Sorts the numerical values in a column in ascending order. Specify the column. Data must be numerical.

Date/time functions to use with LET

CTIME ()

Returns the current time, for example, 9:26:20 AM.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

DATE *("text")*

DATE *(E)*

Returns the date portion corresponding to the argument. The argument should be a text string, such as "3/6/99 10:23", a column, or a stored constant. Text columns and constants should be in one of the Default date/time formats on page 1223.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

ELAPSED *(E)*

Returns the elapsed time, given the start and end times. Enter the column with the end times minus the column with the start times. The columns or values must be in numeric or date/time format. The elapsed time in the output is in minutes and seconds (mm:ss) if the maximum value of the output column is less than one hour; the elapsed time in the output is in hours, minutes, and seconds (hh:mm:ss) if the maximum of the output column is an hour or more.

NETWORKDAYS *(E, E, [E])*

Returns the number of workdays between two dates, inclusive. Specify the start date and end date in the first two arguments. You can enter a column of dates. You can also specify a column of holidays to skip for the optional third argument. If the dates include time portions, Minitab ignores the times. You can also enter single dates in double quotes with the additional use of DATE. For example, to find out the number of workdays between 1/1/09 and 1/31/09, use LET c1 = NETWORKDAYS(date("1/1/09"), date("1/31/09")). Text columns and constants should be in one of the Default date/time formats on page 1223.

NOW *()*

Returns the current date and time, for example, 3/8/2003 9:24.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

TIME *("text")*

TIME *(E)*

Returns the time portion corresponding to the argument. The argument should be a text string, such as "3/6/99 10:23", a column, or a stored constant. Text columns and constants should be in one of the Default date/time formats on page 1223.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

TODAY *()*

Returns today's date, for example, 3/8/2003.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

WHEN *("text")*

WHEN *(E)*

Returns the date and time corresponding to the argument. The argument should be a text string, such as "3/6/99 10:23", or a column or stored constant. Text columns and constants should be in one of the Default date/time formats on page 1223.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

WDAYS *("text", K, ["text"])*

WDAYS *(E, E, [E])*

Offsets the date by the given number of workdays. Enter the start date and the number of workdays to add to the date. Minitab returns the date that is the number of working days (defined as Monday through Friday) from the start date you specify. You can also choose to skip holidays by entering a date or column or dates for the optional third argument. When you enter single dates in double quotes, you need to also use DATE. For example, to find out the date of tenth workday after 1/5/09, use LET c1 = WORKDAYS(date("1/5/09"),10). Text columns and constants should be in one of the Default date/time formats on page 1223.

Note If stored in a column, the result is in a date/time format. If stored in a constant, the result is the numeric representation of the date/time value. Currently, stored constants in Minitab do not have date/time formats.

Logical functions to use with LET

ANY ***(E, K...)***

Returns a 1 if a value equals any value from a set of values, returns a 0 otherwise. Use to flag specified values in a column. In the first argument, specify the column. In the subsequent arguments, specify the values for which to return a 1.

IF ***(C <>=K, "text", ["text"])***

IF ***(C<>=C, C, [C])***

Chooses which of two values to return based on whether a condition is true or false. Conditions can be any numerical or logical expressions. For the first argument, specify the column and the condition using >, <, or =. For the second argument, specify the value to return if the condition is true. The third argument is optional and allows you to specify a value to return if the condition is false. If nothing is specified, Minitab returns a missing value. For example, IF(c1 = 1, "male", "female") returns "male" for values in c1 equal to 1 and returns "female" for the other values in c1.

IF ***(C<>=K,"text", C<>= "text"...["text"])***

IF ***(C<>=K, C, C<>= C...[C])***

Chooses a value to return for each of multiple conditions evaluated sequentially. Conditions can be any numerical or logical expressions. For the first argument, specify the column and the condition using >, <, or =. For the second argument, specify the value to return if the condition is true. You can enter multiple conditions and values. All values must be the same data type (numeric or text). Minitab returns the value corresponding to the first true condition, working from left to right. The final argument is optional and allows you to specify a value to return if all the conditions are false. If nothing is specified, Minitab returns a missing value.

For example, IF(c1 <= 2, "low", c1 <=4, "medium", "high") returns "low" for c1 values less than or equal to 2, "medium" for values less than or equal to 4 but greater than 2, and "high" for the remaining data in c1.

Statistics functions to use with LET

COUNT (C)

Counts the number of nonmissing and missing values in a column.

GMEAN (C)

Calculates the geometric mean, the nth root of the product of the n values. All values must be greater than 0.

MAXIMUM (C)

Calculates the maximum value in a column. Missing values are omitted from the calculation.

MEAN (C)

Calculates the arithmetic mean of all the values in a column. Missing values are omitted from the calculation.

MEDIAN (C)

Calculates the median of the values in a column. Missing values are omitted from the calculation.

MINIMUM (C)

Identifies the minimum value in a column. Missing values are omitted from the calculation.

N (C)

Counts the number of nonmissing values in a column.

NMISS (C)

Counts the number of missing values in a column.

NSCORES (C)
Calculates normal scores.

PERC *(C, E)*
Calculates the sample percentiles For the first argument, enter the column of data. For the second argument, enter the column or constant specifying the desired percentiles, from 0 to 1. Missing values are ignored. For example, to find the first quartile (25th percentile) of a column of data, enter the column number and the percentile 0.25.

RANGE (C)
Calculates the range of the values in a column. Missing values are omitted from the calculation.

SSQ (C)
Calculates the uncorrected sum of squares for a column. Missing values are omitted from the calculations.

STDEV (C)
Calculates the standard deviation for all the values in a column. Missing values are omitted from the calculation.

SUM (C)
Adds all the values in a column. Missing values are omitted from the calculation.

Text functions to use with LET

CLEAN *("text")*

CLEAN *(E)*
Removes all nonprintable characters. Specify the text, column, or stored constant.

CONCATENATE *("text", "text"...)*

CONCATENATE *(E, E...)*
Combines two or more columns or values side-by-side. Specify the text, columns, or stored constants. Minitab treats numbers appearing in a text column (as in a street address or date/time) as text characters and will convert numeric values to text.

FIND *(E, E, [E])*
Identifies the starting position of a string of text within another string of text. For the first argument, specify the text string or the column of text that you want to find. For the second argument, specify the text value or column of text to search. FIND is case-sensitive.

By default, Minitab searches from the first position of each text entry. You can also specify another starting position (the location in the string from where to start the search) by inserting a number for K, the optional third argument.

For example, if c1 contains 234b75, FIND("b7",c1) returns 4, because b7 begins at the 4th position in the text.

Minitab returns a missing value * if the string of text is not found.

FIXED *(E, K, [1])*
Rounds a number to the specified number of decimals and converts it to text with or without commas. For the first argument, specify the number or the column of numeric data you want to convert. For the second argument, specify the number of decimals to retain. If K = 1, Minitab rounds to the nearest tenth. If K = 0 Minitab rounds to the nearest integer. If K = –1, Minitab rounds to a multiple of ten. If you don't enter a second argument, Minitab rounds to 2 decimal places by default.

Minitab inserts commas to separate some place values (hundreds and thousands, hundred-thousands and millions, etc.) in the converted text. If you don't want the commas to appear, enter a value of 1 for the optional third argument.

HTOD *(E)*

Converts hexadecimal values to their equivalent decimal form. The original hexadecimal data must be in text format.

ITEM *("text", E, [E])*

ITEM *E, E, [E])*

Extracts the nth word from a string of text. For the first argument, specify the text or column of text. For the second argument, specify the position of the word to extract.

By default, one or more spaces define where each word begins and ends. If you want to specify other criteria for determining the separation between the words, such as a comma, specify the separator using the optional third argument.

The ITEM function is similar to the WORD function, except that ITEM extracts empty text that occurs between consecutive separators (such as ,,) while WORD ignores the empty string and extracts the text that follows consecutive separators.

LAG *(C, [K])*

Copies the text in the input column to the storage column, moving each value down by the numbers of rows you specify. For the first argument, specify the input column. For the second argument, specify the number of rows the text values should be moved down. By default, if no value is specified, Minitab moves the text values down one row (lag = 1).

LEFT *("text", E)*

LEFT *(C, E)*

Returns the specified number of characters from the beginning of a string of text. For the first argument, specify the text or the column of text values. For the second argument, indicate how many characters from the left you want to retain.

LEN *("text")*

LEN *(C)*

Identifies the number of characters in a string of text. Specify the text or the column of text values.

LOWER *("text")*

LOWER *(C)*

Converts all letters to lowercase. Specify the text or the column of text values.

MID *("text", E, [E])*

MID *(C, E, [E])*

Returns the characters from the middle part of a string of text, given the starting position and the length. For the first argument, specify the text or the column of text values. For the second argument, specify the position of the first character to return.

You can specify an integer for the optional third argument to limit the number of characters that Minitab returns. If you don't enter the third argument, Minitab returns all the characters that follow the starting position.

PAD *("text", E)*

PAD *(C, E)*

Pads text with trailing spaces. For first argument, specify the text or the column of text values. For the second argument, enter the total number of characters needed for the text and the trailing spaces. The padding is indicated by total number of characters minus the number of characters in the text.

PROPER *("text")*

PROPER *(C)*

Capitalizes the first letter in each word and converts all other characters to lowercase. Specify the text or the column of text values.

REPT *("text", E)*

REPT *(C, E)*

Repeats text a given number of times. For the first argument, specify the text or the column of text. For the second argument, specify how many times to repeat the text.

REPLACE *("text", E, E, E)*

REPLACE *(C, E, E, E)*

Replaces a substring of text within a string of text. For the first argument, specify the original text or column of text values. For the second argument, specify the position of the first character to replace. For the third argument, specify how many characters to replace. For the last argument, enter the new text that you want to replace the old text with.

RIGHT *("text", E)*

RIGHT *(C, E)*

Returns the specified number of characters from the end of a string of text. For the first argument, specify the text or the column of text values. For the second argument, indicate how many characters from the right you want to retain.

SEARCH *(E, E, E)*

Identifies the starting position of a string of text within another string of text. For the first argument, specify the text string you want to find. For the second argument, specify the column of text to search. FIND is case-sensitive (distinguishes between b and B).

By default, Minitab searches from the first position of each text entry. You can also specify another starting position (the location in the string from where to start the search) by inserting a number for K, the optional third argument of the function.

SEARCH is similar to FIND, except that SEARCH is not case-sensitive (does not distinguish between b and B).

SUBSTITUTE *("text", E, E, [E])*

SUBSTITUTE *(C, E, E, [E])*

Replaces existing text with new text and allows you to specify which occurrence of the old text you want to replace if the text occurs more than once in a single entry. For the first argument, specify the text or column of text. For the second argument, specify the text that you want to replace. For the third argument, specify the new text you want to substitute. For the last (optional) argument, you can specify which occurrence of the old text you want to replace, if the old text occurs more than once. For example, If c1 contains 600 Pine Lane SUBSTITUTE (c1, "0", "2", 1) returns 620 Pine Lane.

TEXT *(E)*

Converts a numeric value to text. Converting numeric values to text allows you to edit and manipulate the values using text manipulation functions. Specify the number or the column of numeric values.

TRIM *("text")*

TRIM *(C)*

Removes all spaces except single spaces between words. Specify the text or column of text.

WORD *("text", E, [E])*

WORD *(C, E, [E]*

Extracts the n^{th} word from a string of text. For the first argument, specify the text or column of text. For the second argument, specify the position of the word to extract.

By default, one or more spaces define where each word begins and ends. If you want to specify other criteria for determining the separation between the words, such as a comma, specify the separator using the optional third argument.

The WORD function is similar to the ITEM function, except that ITEM extracts the empty text that occurs between consecutive separators (such as the comma and space) while WORD ignores the empty string and extracts the text that follows the consecutive separators.

Trigonometry functions to use with LET

ACSH *(E)*

Calculates the hyperbolic arccosine of a value. Specify the value or the column of values.

ASNH *(E)*

Calculates the hyperbolic arcsine of a value. Specify the value or the column of values.

ATNH *(E)*

Calculates the hyperbolic arctangent of a value. Specify the value or the column of values.

COSH *(E)*

Calculates the hyperbolic cosine of a value. Specify the value or the column of values.

DEGREES *(E)*

Changes radians to degrees. Specify the value or the column of values.

RADIANS *(E)*

Changes degrees to radians. Specify the value or the column of values.

SINH *(E)*

Calculates the hyperbolic sine of a value. Specify the value or the column of values.

TANH *(E)*

Calculates the hyperbolic tangent of a value. Specify the value or the column of values.

Constants functions to use with LET

E ( )
Calculates e = 2.71828...

MISS ( )
Returns the * = missing value symbol in the specified column.

PI ( )
Calculates pi = 3.14159...

Row statistics functions to use with LET

RCOUNT (E, E...)
Counts the number of missing and nonmissing values in the row. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RCOUNT on page 78 to count the number of values in each row while specifying a range of columns, such as C1-C4.

RMAXIMUM (E, E...)
Identifies the maximum value in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RMAXIMUM on page 78 to calculate the maximum value in each row while specifying a range of columns, such as C1-C4.

RMEAN (E, E...)
Calculates the arithmetic mean of the values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RMEAN on page 79 to calculate the mean of each row while specifying a range of columns, such as C1-C4.

RMEDIAN (E, E...)
Identifies the median of the values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RMEDIAN on page 79 to calculate the median of each row while specifying a range of columns, such as C1-C4.

RMINIMUM (E, E...)
Identifies the minimum value in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RMINIMUM on page 79 to calculate the minimum value in each row while specifying a range of columns, such as C1-C4.

RN (E, E...)
Counts the number of nonmissing values in the row. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RN on page 79 to count the number of nonmissing values in each row while specifying a range of columns, such as C1-C4.

RNMISS *(E, E...)*

Counts the number of missing entries in the row. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RNMISS on page 79 to count the number of missing values in each row while specifying a range of columns, such as C1-C4.

RRANGE *(E, E...)*

Calculates the range of values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RRANGE on page 80 to calculate the range of each row while specifying a range of columns, such as C1-C4.

RSSQ *(E, E...)*

Calculates the uncorrected sum of squares of the values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RSSQ on page 80 to calculate the uncorrected sum of squares for each row while specifying a range of columns, such as C1-C4.

RSTDEV *(E, E...)*

Calculates the standard deviation of the values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RSTDEV on page 80 to calculate the standard deviation of each row while specifying a range of columns, such as C1-C4.

RSUM *(E, E...)*

Adds the values in each row. Missing values are omitted from the calculation. You cannot use a hyphen to specify a range of columns. For example, Minitab would interpret C1-C4 as C1 minus C4.

However, you can use the session command RSUM on page 80 to add values in each row while specifying a range of columns, such as C1-C4.

MULTIPLY: Session command for multiplication

MULTIPLY *E...E E*

Note This functionality is available only as a session command, and not in a menu.

Multiplies E by E and stores the result in E. E can be any column, any constant, or any matrix. For columns and constants, MULTIPLY allows up to 50 arguments. If any element of a row is missing, the result is set to missing. If the operation is impossible, the result is also set to missing, *.

Column Statistics

COUNT: Session command for counting the number of values in a column

COUNT C [K]

Counts and optionally stores the number of nonmissing and missing values in a column. (Equivalent to **N total** in the menu command dialog box.)

You can also use this command as a function with LET on page 64.

MEAN: Session command for calculating the arithmetic mean of a column

MEAN

Calculates the arithmetic mean of all the values in a column and optionally stores the result. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64.

MEDIAN: Session command for identifying the median of a column

MEDIAN C [K]

Identifies the median of the values in a column and optionally stores the result. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64.

N: Session command for counting the nonmissing values in a column

N C [K]

Counts and optionally stores the number of nonmissing values in a column. (Equivalent to **N nonmissing** in the menu command dialog box.)

You can also use this command as a function with LET on page 64.

NMISS: Session command for counting the missing values in a column

NMISS C [K]

Counts and optionally stores the number of missing values in a column. (Equivalent to **N missing** in the menu command dialog box.)

You can also use this command as a function with LET on page 64.

RANGE: Session command for calculating a range of values in a column

RANGE C [K]

Calculates the range of values in a column and optionally stores the result. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64.

SSQ: Session command for calculating the uncorrected sum of squares

SSQ C [K]

Calculates the uncorrected sum of squares for a column and optionally stores the result. Missing values are omitted from the calculations.

You can also use this command as a function with LET on page 64.

STDEV: Session command for calculating the standard deviation of all the values in a column

STDEV C [K]

Calculates the standard deviation for all the values in a column and optionally stores the result. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64.

SUM: Session command for adding the values in a column

SUM C [K]

Adds all the values in a column and optionally stores the result. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64.

Row Statistics

RCOUNT: Session command for counting missing and nonmissing values in a row

RCOUNT *E...E* C

Counts and stores the number of missing and nonmissing values in the row. (Equivalent to **N total** in the menu command dialog box.)

Note You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RMAXIMUM: Session command for identifying the maximum value in each row

RMAXIMUM *E...E* C

Identifies and stores the maximum value in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RMINIMUM: Session command for identifying the minimum value in each row

RMINIMUM ***E...E*** **C**

Identifies and stores the minimum value in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RMEAN: Session command for calculating the arithmetic mean in each row

RMEAN

Calculates and stores the arithmetic mean of the values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RMEDIAN: Session command for identifying the median in each row

RMEDIAN

Identifies and stores the median of the values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RN: The session command for counting the nonmissing values in a row

RN ***E...E*** **C**

Counts and stores the number of nonmissing values in the row. (Equivalent to **N nonmissing** in the menu command dialog box.)

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RNMISS: Session command for counting the missing values in a row

RNMISS ***E...E*** **C**

Counts and stores the number of missing entries in the row. (Equivalent to **N missing** in the menu command dialog box.)

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RRANGE: Session command for calculating the range in each row

RRANGE *E...E* **C**

Calculates and stores the range of values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RSSQ: Session command for calculating the uncorrected sum of squares

RSSQ *E...E* **C**

Calculates and stores the uncorrected sum of squares of the values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RSTDEV: Session command for calculating the standard deviation in each row

RSTDEV *E...E* **C**

Calculates and stores the standard deviation of the values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

RSUM: Session command for adding the values in each row

RSUM *E...E* **C**

Adds and stores the values in each row. Missing values are omitted from the calculation.

You can also use this command as a function with LET on page 64. If you use this command as an argument of LET, you cannot use a hyphen to specify a range of values. For example, Minitab would interpret C1-C4 as C1 minus C4.

Standardize

CENTER: Session command for centering data

CENTER C...C, C...C

Centers the data in C...C and stores the results in C...C. Allows you to center and scale columns of data.

When you do not include subcommands, Minitab transforms each input column by subtracting its mean and then dividing by its standard deviation. This is often called standardizing a variable.

LOCATION [K...K]

When you do not include Ks, Minitab transforms each column by subtracting its mean. When you specify one K, Minitab subtracts that value from each column. Otherwise, you must list one K for each column to be centered. Then Minitab subtracts each K from the corresponding column.

SCALE [K...K]

When you use LOCATION, Minitab first subtracts the location. When you do not specify Ks on SCALE, Minitab divides each column by its standard deviation. When you specify one K, Minitab divides each column by K. Otherwise, you must list one K for each column to be centered. Then Minitab divides each column by the corresponding K.

You must specify K as greater than 0.

MINMAX [K K]

When you do not specify Ks, Minitab transforms all columns (linearly) to have minimum –1 and maximum +1. When you specify both Ks, Minitab transforms all columns (linearly) to have the first K as minimum and the second K as maximum.

Make Patterned Data

SET: Session command for entering data into a column

SET C

Inputs data, from the keyboard or from an ASCII file, into one column.

If there is any data in the column you list, SET erases that data and replaces it with the new data you enter. SET is especially useful for entering data that follow a pattern, such as the numbers 1 through 10, into a column. For more information, go to Entering patterned data for the SET session command on page 1226.

If you execute SET from the menu, the FORMAT and NOBS options are not available, the form of the data is restricted, and data cannot be read from a file.

As a simple example, the following command language puts the numbers 2, 7, 9, 3.8, and 22 in column C7.

```
SET C7
 2  7 9
 3.8 22
END
```

FILE *'filename'*

Enters data from the specified ASCII file. If the file has an extension other than DAT and/or if it is not in your current directory, include the file name extension and the path within the quotation marks.

FORMAT

The FORMAT subcommand for SET is similar to that for READ (for information, go to READ data into a matrix on page 44 or READ data into columns on page 42). For example, the following command language reads 10 numbers from each data line-the first from spaces 1 and 2, the second from spaces 3 and 4, and so on.

```
SET C1;
  FILE "MYFILE";
  FORMAT (10F2).
```

Text data can be input using the FORMAT subcommand. The field width can be up to 80 characters. You cannot mix text and numeric data in a single column.

Here is an example that inputs four "words" using A format into C1.

```
SET C1;
 FORMAT (A11, T16, A2).
 ABCDEFGHIJK   LM
 abcdefghijk   lm
END
```

For more information, go to Valid format items on page 1269.

NOBS K

Specifies the number of observations (rows) to be read. If an END on page 39 subcommand or end-of-file is encountered before K observations are read, NOBS is ignored. NOBS is useful when you want to read just the first portion of a file. NOBS is also useful when you prompt the user for information. For more information, go to Prompting a user for information on page 1260.

SKIP K

Tells Minitab to skip K lines at the top of the data file before beginning to READ or SET data into the file. This is most useful when you have one or more lines of text, such as column names and titles, at the top of a data file that you want to import into Minitab.

DECIMAL "." or ","

Specifies a period or comma as the decimal separator.

TSET: Session command for creating data that follow complicated patterns

TSET

Using TSET, you can create data that follow more complicated patterns. You can also use TSET to import text files.

TSET is the same as SET except there is no through construct, and you list text constants, stated or stored, instead of numbers. Repeat factors are integers, stated or stored. A text string must be enclosed in double quotes as in LET.

As in SET, you can input data from a file in addition to typing it from the keyboard.

FILE *"filename"*

Enters data from the specified file. If the file has an extension other than DAT and/or if it is not in your current directory, include the file name extension and the path within the quotation marks.

Examples

This example contains text constants and repeat factors.

```
TSET C1
 5("Red" "Green")  2("Yellow")  3("Green")
 END
LET K1 = "Low"
LET K2 = 10
TSET   C1
  5(K1 "Medium" "High")K2  K1  K1
  END
```

DSET: Session command for making patterned data

DSET C

Creates a new column of dates and/or times that follow a pattern.

For example, the commands store the dates 1/1/15, 1/2/15, ..., 1/31/15 into C1.

```
DSET C1;
 DSTART "1/1/15";
 DEND "1/31/15";
 DAY 1.
```

Note Minitab's default date/time formats can change depending on the Regional Settings of your device. For more information, go to Default date/time formats on page 1223.

LIST

Date/time K...K or "text" ..."text".

FORMAT K

The format statement can be used for two purposes. It provides the format for the resulting patterned data. If no format is given then the resulting column will be in the format of the last date/time value given either on the DSTART or DEND subcommand lines. Also, it can be utilized as a user-defined format for specifying the input format of the start and end date/time value. If no format statement is given, the date/time values on the DSTART and DEND subcommands must be in a default format. For example:

```
DSET C1;
DSTART "1/1/15";
DEND "1/31/15";
FORMAT (DTm-d-yy);
DAY 1.
```

The format statement must begin with the characters DT followed by the desired date/time format. The preceding example displays dates in the form 1-1-15.

Note Minitab's default date/time formats can change depending on the Regional Settings of your device. For more information, go to Default date/time formats on page 1223.

DSTART K

DSTART specifies the beginning date/time value. The argument can be either a stored text constant or a text string in double quotation marks. If you want to use a non-default date/time format for the text given with DSTART and DEND, you must use the FORMAT subcommand. (If you are using a default date/time format, FORMAT is not necessary.)

Note Minitab's default date/time formats can change depending on the Regional Settings of your device. For more information, go to Default date/time formats on page 1223.

DEND K

DEND specifies the ending date/time value. The argument can be either a stored text constant or a text string in double quotation marks. If you want to use a non-default date/time format for the text given with DSTART and DEND, you must use the FORMAT subcommand. (If you are using a default date/time format, FORMAT is not necessary.)

Note Minitab's default date/time formats can change depending on the Regional Settings of your device. For more information, go to Default date/time formats on page 1223.

RLIST K

RLIST specifies the number of times to repeat the list. K must be a positive integer greater than or equal to 1.

RVALUE K

RVALUE specifies the number of times to repeat each value. K must be a positive integer greater than or equal to 1.

Subcommands for incrementing the date/time data

You must use one (and only one) of the following subcommands with DSET to indicate how to increment the date/time data. K must be a positive integer greater than or equal to 1.

The patterned data ends when a date/time value is greater than or equal to the end date/time value (DEND). The end date/time value is part of the column only if the pattern equals the end date/time value.

Note Minitab's default date/time formats can change depending on the Regional Settings of your device. For more information, go to Default date/time formats on page 1223.

DAY

Increments day by K units.

WDAY

Increments workday (M–F) by K units.

WEEK

Increments week by K units.

MONTH

Increments month by K units.

QUARTER

Increments quarter by K units.

YEAR

Increments year by K units.

HOUR

Increments hour by K units.

MINUTE

Increments minute by K units.

SECOND

Increment second by K units.

HUNDREDTH
Increment hundredth by K units.

THOUSANDTHS
Increment thousandths by K units.

Make Mesh Data

MESH: Session command for making mesh data

MESH C C
Creates a mesh of regular x-y data and stores the data in C and C. (Data with a regular shape form a grid with evenly spaced intervals. Data with an irregular shape are not located on evenly spaced intervals.)

You can use the mesh data for drawing surface and wireframe plots, with an option to create the z-data at the same time.

XMESH K K K (optional)
Specifies the lowest x-value, the highest x-value, and the number of x-values. The default is – 5 5 11.

YMESH K K K (optional)
Specifies the lowest y-value, the highest y-value, and the number of y-values. The default is – 5 5 11.

%USERFUNC K C C C (optional)
Specifies your own function as indicated by function number K in the macro %USERFUNC. For instructions to add your own function, go to Add your own function on page 1214.

PARAMS K...K (optional)
Specifies parameters for the functions. See the table below for formulas and the number of parameters.

Functions for z-values (optional)

The function subcommands in the following table generate Z-values for the mesh and store them in C.

Subcommand	Equation	Parameters (in order)
BVNORMAL C C	$z = ae^{(-b/2(1-\rho^2))}$	μ_1 = MU1, σ_1 = S1, μ_2 = MU2, σ_2 = S2, ρ = RHO ($\sigma_1 \neq 0$, $\sigma_2 \neq 0$, $-1 < \rho < 1$) The following variables are derived from the above variables, as follows: $a = \frac{1}{2 \times \pi \times \sigma_1 \times \sigma_2 \times \sqrt{1-\rho^2}}$ $b = \left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2 \times \rho \times \frac{x-\mu_1}{\sigma_1} \times \frac{y-\mu_2}{\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2$
BOWL C	$z = \frac{\left(\frac{x}{A}\right)^2 + \left(\frac{y}{B}\right)^2}{C}$	A, B, C ($A \neq 0$, $B \neq 0$, $C \neq 0$)

Subcommand	Equation	Parameters (in order)
CONE C	$z = C\sqrt{\left(\frac{x}{A}\right)^2 + \left(\frac{y}{B}\right)^2}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)
COWBOYHAT C	$z = \frac{\sin\sqrt{\left(\frac{x}{A}\right)^2 + \left(\frac{y}{B}\right)^2}}{C}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)
EGGCARTON C	$z = \frac{\sin\frac{x}{A} \times \cos\frac{x}{A} + \sin\frac{y}{B} \times \cos\frac{y}{B}}{C}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)
HEMISPHERE C	$z = C\left(\sqrt{R^2\left(\left(\frac{x}{A}\right)^2 + \left(\frac{y}{B}\right)^2\right)}\right)$ where if z < 0, set z = 0	A, B, C, R (A ≠ 0, B ≠ 0, C ≠ 0, R ≠ 0)
HILLANDDALE C	$z = \frac{\sin\frac{x}{A} + \cos\frac{y}{B}}{C}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)
SADDLE C	$z = \frac{\left(\frac{x}{A}\right)^2 - \left(\frac{y}{B}\right)^2}{C}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)
WAVE C	$z = \frac{\sin\left(\frac{x}{A} + \frac{y}{B}\right)}{C}$	A, B, C (A ≠ 0, B ≠ 0, C ≠ 0)

Make Indicator Variables

INDICATOR: Session command for creating indicator variables

INDICATOR C C...C

Creates indicator variables (also called dummy variables) that you can use in a regression analysis. If you use REGRESS on page 170, you do not need to create indicator variables.

The smallest number in C2 is 2 and the largest is 6. INDICATOR creates one indicator variable for each unique value.

- C11 is the indicator variable for the value 2. C11 contains a 1 in every row where C2 contains a 2, and 0 otherwise.
- C12 is the indicator variable for the value 3. C12 contains a 1 in every row where C2 contains a 3, and 0 otherwise.
- C13 is the indicator variable for 5. C13 contains a 1 in every row where C2 contains 5, and 0 otherwise.
- C14 is the indicator variable for 6. C13 contains a 1 in every row where C2 contains 6, and 0 otherwise.

If C2 contains an * (missing data code), then all indicator variables are also set to *.

The number of storage columns must be equal to the number of distinct values (not including *) in the input column. Up to 100 storage columns are allowed on INDICATOR.

The following command language is an example of using INDICATOR.

```
INDICATOR  C2, C11-C14
```

Before using INDICATOR

C1
2
2
5
3
6
5
*

After using INDICATOR

C1	...	C11	C12	C13	C14
2	...	1	0	0	0
2	...	1	0	0	0
5	...	0	0	1	0
3	...	0	1	0	0
6	...	0	0	0	1
5	...	0	0	1	0
*	...	*	*	*	*

Random Data

BASE: Session command for fixing a starting number for the random number generator

BASE K

Fixes a starting point for Minitab's random number generator.

Minitab has a long string of random numbers available. Minitab normally chooses its own starting point for this process. If Minitab always started at the beginning of the list, you would always get the same data. To avoid this, Minitab uses the time of day to choose a random starting point in the string.

However, you might want to control where Minitab starts the string. For example, you may wish to repeat a sequence by generating the same set of random data. In this case, the BASE command tells the random number generator where to start. The generator will use this base until you set a new BASE or exit Minitab.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

RANDOM: Session command for generating random data

RANDOM *E [E]*

Use RANDOM to generate and store a random sample of one or more observations from a specified distribution.

The subcommand specifies the distribution. If no subcommand is given, data are simulated from a normal distribution with mu = 0 and sigma = 1.

Use the BASE on page 87 command to generate the same set of data more than once.

Suppose you want 50 random samples, each containing 20 observations, from a binomial distribution with number of trials n = 5, and probability of success p = 0.3. To place each sample in a separate column, type the following commands.

```
RANDOM 20 C1-C50;
BINOMIAL 5 0.3.
```

CHISQUARE K

Specifies distribution to sample, with degrees of freedom = K.

NORMAL [K [K]]

Specifies distribution to sample. Generates data from a standard normal. Optionally, specify mean = K, and standard deviation = K.

MNORMAL C *M*

Specifies distribution to sample, with mean column = C, and variance-covariance matrix = M.

F K K

Specifies distribution to sample, with numerator degrees of freedom = K, denominator degrees of freedom = K.

T K

Specifies distribution to sample, with degrees of freedom = K.

UNIFORM [K K]

Specifies distribution to sample. Generates data using lower endpoint = 0.0 and upper endpoint = 1.0. Optionally, specify lower endpoint = K and upper endpoint = K.

BERNOULLI K

Specifies distribution to sample, with probability of success = K.

BINOMIAL K K

Specifies distribution to sample, with number of trials = K and event probability = K.

GEOMETRIC K

Specifies distribution to sample, with event probability = K.

NONEVENT

Models the number of nonevents before the first event occurs.

TOTAL

Models the total number of trials needed to produce one event.

NEGBINOMIAL K K

Specifies distribution to sample, with event probability = K and number of events needed = K.

NONEVENT
Models the number of nonevents before the specified number of events occurs.

TOTAL
Models the total number of trials needed to produce the specified number of events.

HYPERGEOMETRIC K K K
Specifies distribution to sample, with population size = K, event count in population = K, and sample size = K.

DISCRETE C C
Specifies distribution to sample, with values in C and probabilities in C.

INTEGER K K
Specifies distribution to sample, with discrete uniform on integers from minimum value = K to maximum value = K.

POISSON K
Specifies distribution to sample, with mean = K.

BETA K K
Specifies distribution to sample, with first shape parameter = K and second shape parameter = K.

CAUCHY [K [K]]
Specifies distribution to sample. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

EXPONENTIAL [K [K]]
Specifies distribution to sample. Generates data using mean = 1.0 and threshold = 0.0. Optionally, specify mean = K and threshold = K.

GAMMA K K [K]
Specifies distribution to sample, with shape = K, scale = K, and optionally, threshold = K.

LAPLACE [K [K]]
Specifies distribution to sample. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LOGISTIC [K [K]]
Specifies distribution to sample. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LLOGISTIC [K [K [K]]]
Specifies distribution to sample. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

LNORMAL [K [K [K]]]
Specifies distribution to sample. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

SEXTREME [K [K]]
Specifies distribution to sample. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

TRIANGULAR K K K
Specifies distribution to sample, with lower endpoint = K, mode = K, and upper endpoint = K.

WEIBULL K K [K]
Specifies distribution sample, with shape = K, scale = K, and optionally, threshold = K.

SAMPLE: Session command for generating rows of random data from specified columns

SAMPLE K C...C
Generates K rows of random data from specified input columns, C...C, and stores in specified storage columns, C...C.

Takes a random sample of rows. If you use REPLACE, you can select the same row more than once. If you use NOREPLACE, you cannot select the same row more than once. If you do not use a subcommand, Minitab samples without replacement.

Tip If there are K rows in C...C columns, you can randomize their order by sampling all K rows and storing in the original input columns. For example, `SAMPLE K C1 C1`.

REPLACE
When you sample with replacement, a selected observation goes back into the pool of possible choices and can be selected again.

NOREPLACE
When you sample without replacement, a selected observation can be chosen only once.

Probability Distributions

PDF: Session command for calculating the probability distribution of a continuous random variable

PDF *E [E]*
Calculates density values or probabilities for the specified values in E from a standard normal distribution or another specified distribution and stores in E.

- For a discrete distribution, the probability distribution function (pdf) calculates probabilities for the specified values (sometimes called the discrete probability distribution function). If you specify a discrete distribution (BINOMIAL, GEOMETRIC, NEGBINOMIAL, HYPERGEOMETRIC, DISCRETE, INTEGER, POISSON), the arguments on the PDF line are optional. If you do not specify arguments, Minitab displays a table of the distribution. If you execute PDF from the menu, you must supply the input columns.
- For a continuous distribution, pdf calculates the continuous probability density function (often called the density function).
- If you do not specify a distribution, results are generated for a normal distribution with mu = 0 and sigma = 1.

Storage is optional. If you specify a storage column, pdf values are stored there and are not displayed. If you do not specify a storage column, Minitab displays pdf values.

CHISQUARE K [K]
Specifies distribution with degrees of freedom = K.

NORMAL [K [K]]
Specifies distribution. Generates data from a standard normal; optionally specify mean = K, and standard deviation = K.

F K K
Specifies distribution, with numerator degrees of freedom = K, denominator degrees of freedom = K.

T K
Specifies distribution, with degrees of freedom = K.

UNIFORM [K K]
Specifies distribution. Generates data using lower endpoint = 0.0 and upper endpoint = 1.0. Optionally, specify lower endpoint = K and upper endpoint = K.

BINOMIAL K K
Specifies distribution, with number of trials = K and event probability = K.

GEOMETRIC K
Specifies distribution, with event probability = K.

NONEVENT
Models the number of nonevents before the first event occurs.

TOTAL
Models the total number of trials needed to produce one event.

NEGBINOMIAL K K
Specifies distribution, with event probability = K and number of events needed = K.

NONEVENT
Models the number of nonevents before the specified number of events occurs.

TOTAL
Models the total number of trials needed to produce the specified number of events.

HYPERGEOMETRIC K K K
Specifies distribution, with population size = K, event count in population = K, and sample size = K.

DISCRETE C C
Specifies distribution, with values in C and probabilities in C.

INTEGER K K
Specifies distribution, with discrete uniform on integers from minimum value = K to maximum value = K.

POISSON K
Specifies distribution, with mean = K.

BETA K K
Specifies distribution, with first shape parameter = K and second shape parameter = K.

CAUCHY [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

EXPONENTIAL [K [K]]
Specifies distribution. Generates data using mean = 1.0 and threshold = 0.0. Optionally, specify mean = K and threshold = K.

GAMMA K K [K]
Specifies distribution, with shape = K, scale = K, and optionally, threshold = K.

LAPLACE [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LOGISTIC [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LLOGISTIC [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

LNORMAL [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

SEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

TRIANGULAR K K K
Specifies distribution, with lower endpoint = K, mode = K, and upper endpoint = K.

WEIBULL K K [K]
Specifies distribution, with shape = K, scale = K, and optionally, threshold = K.

CDF: Session command for calculating the cumulative probability of an x-value

CDF *E [E]*
Calculates probabilities for the specified values in E from a standard normal distribution or another specified distribution and stores the probabilities in E.

The cumulative distribution function, cdf, for any value x is the probability that a random variable with the specified distribution has a value less than or equal to x. That is:

CDF (x) = Pr (X < x)

- If you specify a discrete distribution (BINOMIAL, GEOMETRIC, NEGBINOMIAL, HYPERGEOMETRIC, DISCRETE, INTEGER, POISSON), then the arguments on the CDF line are optional. If you do not specify an argument, then Minitab displays a table of the distribution. (Entries where the cdf is less than 0.00005 or greater than 0.99995 might not be displayed in this table.) If you execute CDF from the menu, you must supply the input columns.
- If you do not specify a distribution, then the results are generated for a normal distribution with mu = 0 and sigma = 1.

Storage is optional. If you specify a storage column, the cdf values are stored there and are not displayed. If you do not specify a storage column, Minitab displays the cdf values.

CHISQUARE K [K]
Specifies distribution with degrees of freedom = K, and, optionally, noncentrality parameter = K.

NORMAL [K [K]]
Specifies distribution. Generates data from a standard normal. Optionally, specify mean = K and standard deviation = K.

F K K [K]
Specifies distribution, with numerator degrees of freedom = K, denominator degrees of freedom = K, and, optionally, noncentrality parameter = K.

T K [K]
Specifies distribution, with degrees of freedom = K and, optionally, noncentrality parameter = K.

UNIFORM [K K]
Specifies distribution. Generates data using lower endpoint = 0.0 and upper endpoint = 1.0. Optionally, specify lower endpoint = K and upper endpoint = K.

BINOMIAL K K
Specifies distribution, with number of trials = K and event probability = K.

GEOMETRIC K
Specifies distribution, with event probability = K.

NONEVENT
Models the number of nonevents before the first event occurs.

TOTAL
Models the total number of trials that are needed to produce one event.

NEGBINOMIAL K K
Specifies distribution, with event probability = K and number of events that are needed = K.

NONEVENT
Models the number of nonevents before the specified number of events occurs.

TOTAL
Models the total number of trials that are needed to produce the specified number of events.

HYPERGEOMETRIC K K K
Specifies distribution, with population size = K, event count in population = K, and sample size = K.

DISCRETE C C
Specifies distribution, with values in C and probabilities in C.

INTEGER K K
Specifies distribution, with discrete uniform on integers from minimum value = K to maximum value = K.

POISSON K
Specifies distribution, with mean = K.

BETA K K
Specifies distribution, with first shape parameter = K and second shape parameter = K.

CAUCHY [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

EXPONENTIAL [K [K]]
Specifies distribution. Generates data using mean = 1.0 and threshold = 0.0. Optionally, specify mean = K and threshold = K.

GAMMA K K [K]
Specifies distribution, with shape = K, scale = K, and optionally, threshold = K.

LAPLACE [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LOGISTIC [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LLOGISTIC [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

LNORMAL [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

SEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

TRIANGULAR K K K
Specifies distribution, with lower endpoint = K, mode = K, and upper endpoint = K.

WEIBULL K K [K]
Specifies distribution, with shape = K, scale = K, and optionally, threshold = K.

INVCDF: Session command for calculating the variable for a cumulative probability

INVCDF *E [E]*

Calculates the inverse of the cdf for the specified values in E from a standard normal distribution or another specified distribution and stores in E.

INVCDF returns the inverse of the cumulative distribution function (cdf), meaning that for a given probability, p, INVCDF finds a value of x such that p = CDF(x).

For discrete distributions, INVCDF (p) = min {x such that CDF(x) > p}.

If you do not specify a distribution, results are generated for a normal distribution with mu = 0 and sigma = 1.

Storage is optional. If you specify a storage column, the INVCDF values are stored there and are not displayed. If you do not specify a storage column, Minitab displays the INVCDF values.

CHISQUARE K [K]

Specifies distribution with degrees of freedom = K, and, optionally, noncentrality parameter = K.

NORMAL [K [K]]

Specifies distribution. Generates data from a standard normal distribution. Optionally, specify mean = K, and standard deviation = K.

F K K [K]

Specifies distribution, with numerator degrees of freedom = K, denominator degrees of freedom = K, and, optionally, noncentrality parameter = K.

T K [K]

Specifies distribution, with degrees of freedom = K and, optionally, noncentrality parameter = K.

UNIFORM [K K]

Specifies distribution. Generates data using lower endpoint = 0.0 and upper endpoint = 1.0. Optionally, specify lower endpoint = K and upper endpoint = K.

BINOMIAL K K

Specifies distribution, with number of trials = K and event probability = K.

GEOMETRIC K

Specifies distribution, with event probability = K.

NONEVENT

Models the number of nonevents before the first event occurs.

TOTAL

Models the total number of trials that are needed to produce one event.

NEGBINOMIAL K K

Specifies distribution, with event probability = K and number of events needed = K.

NONEVENT

Models the number of nonevents before the specified number of events occurs.

TOTAL

Models the total number of trials that are needed to produce the specified number of events.

HYPERGEOMETRIC K K K
Specifies distribution, with population size = K, event count in population = K, and sample size = K.

DISCRETE C C
Specifies distribution, with values in C and probabilities in C.

INTEGER K K
Specifies distribution, with discrete uniform on integers from minimum value = K to maximum value = K.

POISSON K
Specifies distribution, with mean = K.

BETA K K
Specifies distribution, with first shape parameter = K and second shape parameter = K.

CAUCHY [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

EXPONENTIAL [K [K]]
Specifies distribution. Generates data using mean = 1.0 and threshold = 0.0. Optionally, specify mean = K and threshold = K.

GAMMA K K [K]
Specifies distribution, with shape = K, scale = K, and, optionally, threshold = K.

LAPLACE [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LOGISTIC [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

LLOGISTIC [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

LNORMAL [K [K [K]]]
Specifies distribution. Generates data using location = 0.0, scale = 1.0, and threshold = 0.0. Optionally, specify location = K, scale = K, and threshold = K.

SEXTREME [K [K]]
Specifies distribution. Generates data using location = 0.0 and scale = 1.0. Optionally, specify location = K and scale = K.

TRIANGULAR K K K
Specifies distribution, with lower endpoint = K, mode = K, and upper endpoint = K.

WEIBULL K K [K]
Specifies distribution, with shape = K, scale = K, and optionally, threshold = K.

Resampling Analyses

BTFT: Session command for calculating a 1-sample bootstrap confidence interval of a function

BTFT C K

BTFT calculates a confidence interval for a function in a population. The C specifies the column that contains a sample from the population. The K specifies the number of resamples. The number of resamples can be from 1 to 10,000. The available functions are mean, median, sum, variance, and standard deviation.

Use the appropriate subcommand to choose the function. Specify only one of MEAN, MEDIAN, SUM, VARIANCE, and STDEV.

MEAN

Calculates a confidence interval for the mean.

MEDIAN

Calculates a confidence interval for the median.

SUM

Calculates a confidence interval for the sum.

VARIANCE

Calculates a confidence interval for the variance.

STDEV

Calculates a confidence interval for the standard deviation.

Options

ITYPE K

Enter K to specify the type of confidence interval.

Value of K	Format of the alternative hypothesis
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SSAMPLE C

Stores the statistics from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TSTATISTICS

Displays summary statistics for the sample: mean, standard deviation, variance, sum, minimum, median, and maximum.

TBOOTSTRAP

Displays the confidence interval and the associated statistics.

GHISTOGRAM

Displays a histogram of the resampled statistics. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GINDIVIDUAL

Displays an individual value plot that compares the resample to the original sample. GINDIVIDUAL does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

BTPR: Session command for calculating a 1-sample bootstrap confidence interval of a proportion

BTPR C K

BTPR K K K

BTPR calculates a confidence interval for a proportion in a population. In the first parameterization, the C specifies the column that contains a sample from the population. The column must contain 1 or 2 unique, non-missing values. The K specifies the number of resamples. The number of resamples can be from 1 to 10,000.

In the second parameterization, the first K represents the number of events. The first K must be greater than or equal to 1. The second K represents the number of trials. The number of trials must be greater than the number of events and less than 100,000. The third K is the number of resamples. The number of resamples must be from 1 to 10,000.

EVENT K

When the sample data are in a column, use EVENT to specify the value in the column that the proportion should be for. The value of K matches one of the values in the sample column. Enclose text values in quotation marks. For example, to calculate an interval for the proportion of a "Fail" in the sample column, enter EVENT "Fail".

Options

ITYPE K

Enter K to specify the type of confidence interval.

Value of K	Format of the alternative hypothesis
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SSAMPLE C

Stores the proportions from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TSTATISTICS

Displays summary statistics for the sample: sample size and proportion.

TBOOTSTRAP

Displays the confidence interval and the associated statistics.

GHISTOGRAM

Displays a histogram of the resampled proportions. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GBARCHART

Displays a bar chart that compares the resample to the original sample. GBARCHART does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

BTTM: Session command for calculating a 2-sample bootstrap confidence interval for the difference of means

BTTM C C K

BTTM calculates a confidence interval for the difference between the means of two populations. The columns can have one of two formats:

- The first column contains the measurements. The second column identifies which population each observation is from.
- Each column contains a sample from a different population. Use UNSTACKED to specify that the data are in this format.

The value of K is the number of resamples. The number of resamples must be from 1 to 10,000.

The difference is the first sample minus the second sample. When the data are stacked, the first sample depends on the values in the identification column. The first value is the lowest number, the first value alphabetically, the earliest date, or the earliest value in a previously-specified value order. When the data are unstacked, the first sample is in the first C after BTTM.

UNSTACKED

Indicates that the columns after BTTM contain different samples.

Options

ITYPE K

Enter K to specify the type of confidence interval.

Value of K	Format of the alternative hypothesis
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SDIFFERENCES C

Stores the differences from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TOBSERVED

Displays summary statistics for the samples: mean, standard deviation, variance, sum, minimum, median, and maximum.

TDIFFERENCES

Displays the difference between the sample means.

TBOOTSTRAP

Displays the confidence interval and the associated statistics.

GHISTOGRAM

Displays a histogram of the resampled differences. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GINDIVIDUAL

Displays an individual value plot that compares the resamples to the original samples. GINDIVIDUAL does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

RNMN: Session command for performing a 1-sample randomization test of a mean

RNMN C K K

RNMN performs a hypothesis test that the mean of the population equals the second K. The C specifies the column that contains a sample from the population. The first K specifies the number of resamples. The number of resamples can be from 1 to 10,000.

Options

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SMEAN C

Stores the means from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TOBSERVED

Displays summary statistics for the sample: mean, standard deviation, variance, sum, minimum, median, and maximum.

TRANDOMIZATION

Displays the results of the hypothesis test.

GHISTOGRAM

Displays a histogram of the resampled means. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GINDIVIDUAL

Displays an individual value plot that compares the resample to the original sample. GINDIVIDUAL does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

RNPR: Session command for performing a 1-sample randomization test of a proportion

RNPR C K K

RNPR K K K K

RNPR performs a hypothesis test on a proportion in a population. In the first parameterization, the C specifies the column that contains a sample from the population. The column must contain 1 or 2 unique, non-missing values. The first K specifies the number of resamples. The number of resamples can be from 1 to 10,000. The last K is the hypothesized proportion. The hypothesized proportion must be a value between 0 and 1.

In the second parameterization, the first K represents the number of events. The first K must be greater than or equal to 1. The second K represents the number of trials. The number of trials must be greater than the number of events and less than 1 million. The third K is the number of resamples. The number of resamples must be from 1 to 10,000. The last K is the hypothesized proportion. The hypothesized proportion must be a value between 0 and 1.

Options

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SPROPORTION C

Stores the proportions from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TOBSERVED

Displays summary statistics for the sample: sample size and proportion.

TRANDOMIZATION

Displays the results of the hypothesis test.

GHISTOGRAM

Displays a histogram of the resampled proportions. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GBARCHART

Displays a bar chart that compares the resample to the original sample. GBARCHART does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

RNTM: Session command for performing a 2-sample randomization test of means

RNTM C C K

RNTM performs a hypothesis test that the means of two populations are equal. The columns can have one of two formats:

- The first column contains the measurements. The second column identifies which population each observation is from.
- Each column contains a sample from a different population. Use UNSTACKED to specify that the data are in this format.

The value of K specifies the number of resamples. The number of resamples must be from 1 to 10,000.

The difference is the first sample minus the second sample. When the data are stacked, the first sample depends on the values in the identification column. The first value is the lowest number, the first value alphabetically, the earliest date, or the earliest value in a previously-specified value order. When the data are unstacked, the first sample is in the first C after RNTM.

UNSTACKED

Indicates that the columns after RNTM contain different samples.

Options

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	The mean of the first sample is less than the mean of the second sample.
0 (default)	The mean of the first sample is not equal to the mean of the second sample.
1	The mean of the first sample is greater than the mean of the second sample.

BASE K

Fixes a starting point for Minitab's random number generator.

Normally, Minitab uses the time of day to start the random number generator. Instead, you can use the BASE subcommand to specify a number to use to start the generator. Use the BASE subcommand so that you can generate the same sample repeatedly.

Note If you use the same base on different platforms or different versions of Minitab, you might not get the same random number sequence.

SDIFFERENCES C

Stores the differences from the resampling process.

Results

Use the following subcommands to control what output Minitab produces.

TOBSERVED

Displays summary statistics for the samples: mean, standard deviation, variance, sum, minimum, median, and maximum.

TRANDOMIZATION

Displays the results of the hypothesis test.

TDIFFERENCES

Displays the difference between the sample means.

GHISTOGRAM

Displays a histogram of the resampled differences. GHISTOGRAM does not work if the number of resamples is 1 or if you specify GINDIVIDUAL.

GINDIVIDUAL

Displays an individual value plot that compares the resamples to the original samples. GINDIVIDUAL does not work if the number of resamples is greater than 1 or if you specify GHISTOGRAM.

Matrices

READ data into a matrix

READ K K M

Puts numbers into a matrix. To input data to columns, go to READ data into columns on page 42. You can specify the filename as either the name of the file in double quotes, or a stored text constant. If the file has an extension other than DAT and/or if it is not in your current directory, include the file name extension and the path within quotation marks.

You can use either spaces or commas to separate the data in the matrix.

You must specify the dimension of the matrix in the READ command. The first K gives the number of rows, the second K the number of columns. The M is the matrix identifier for storage. If a file name is not used, READ is followed by data lines, each containing one row of the matrix. The following command creates the following matrix.

Command

```
READ  3  4  M2
1  2  3  4
5  6  7  8
9  10  11  12
END
```

Matrix

```
1  2  3  4
5  6  7  8
9  10  11  12
```

FILE *"filename"*

FILE K

Reads or inserts data from the specified text file.

DEFINE: Session command for defining a constant matrix

DEFINE K K K *M*

Defines a constant matrix where K is the value, K is the number of rows, K is the number of columns, and M is the stored matrix. For example, the following command language creates a matrix, M1, has 4 rows and 3 columns, and all of whose entries are 1.

```
DEFINE 1 4 3 M1
```

DIAGONAL: Session command for creating a matrix from a column

DIAGONAL C *M*

Forms a matrix out of a column. If C has n entries, then M will be an n x n matrix with C as its diagonal and zeros elsewhere.

DIAGONAL *M* C

Takes the diagonal of a matrix and puts it into a column.

INVERT: Session command for replacing a matrix value with its inverse

INVERT *M M*

Replaces each value in the matrix with its inverse. The first M is the matrix to be replaced. The second M is the new inverted matrix. The matrix must be square.

TRANSPOSE: Session command for changing rows to columns, and columns to rows

TRANSPOSE C...C

TRANSPOSE C *M*

TRANSPOSE *M* C

TRANSPOSE *M [M]*

TRANSPOSE reconfigures data so that rows become columns and columns become rows.

- TRANSPOSE C...C transposes the values in the columns.
- TRANSPOSE C M transposes the values in a column to a matrix.
- TRANSPOSE MC transposes the values in a single row matrix to a column.
- TRANSPOSE M [M] transposes the values in a matrix to another matrix.

VARNAMES C

Specifies a column that contains variable names for the transposed columns.

STORE C...C

STORE *M*

Specifies where the transposed data are to be written. You can store the data in columns or a matrix. When you transpose data to columns, the number of columns specified must equal the number of rows in the original columns.

If you do not use STORE, the data are stored in a new worksheet.

LABELS C

Specifies the column in which the labels of the transposed column or columns or matrix are stored.

NEWWS

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in "text" or in a constant (K) argument, then Minitab uses the default naming of worksheets.

Important You cannot use NEWWS in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

AFTER

Appends the transposed data into empty columns after the last column of data in the current worksheet.

Important You cannot use AFTER in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

EIGEN: Session command for calculating eigenvalues

EIGEN ***M*** **C** ***[M]***

Calculates eigenvalues (also called characteristic values or latent roots) and eigenvectors for a symmetric matrix.

The eigenvalues are stored in decreasing order of magnitude down the column. The eigenvectors are stored as columns of the matrix. The first column corresponds to the first eigenvalue (largest magnitude), the second column to the second eigenvalue, and so on.

Data

INFO: Session command for summarizing the current worksheet

INFO [C...C]

Summarizes the current worksheet.

If no columns are specified, INFO prints a list of all columns used with their names and counts, all stored constants, all matrices. If there are missing observations, a count of these is also given. If a column contains text data, the letter T is printed to the left of the column. If columns have assigned formulas, these are printed along with the method selected for updating the calculations (manual or automatic). If you list columns, information is given on just those columns.

MERGE: Session command for merging two worksheets into one worksheet

MERGE

Merges two worksheets into one new worksheet. You can combine any two open worksheets using MERGE. Unlike the merging option available when opening additional worksheets, MERGE duplicates, and then combines, the information from two original worksheets into a new worksheet. The default setting combines the worksheets, side by side and with their existing attributes, into the new worksheet, named Merge Worksheet by default.

You can also customize the merged worksheet using the optional subcommands. You can use BY to combine the worksheets according to the order and length of one or more columns. In addition, you can specify whether or not to include unmatched values, missing values or multiple values from one or both of the BY columns. You can also specify which columns will be included from each original worksheet using INCLUDE.

Note Stored constants, matrices, DOE objects, and worksheet descriptions do not transfer into the merged worksheet.

NAME K

Specifies the name K for the new worksheet. The argument is a text constant, stated or stored. By default, the name is *Merge Worksheet*.

WORKSHEET K

Specifies the name of one of the worksheets to be merged. The argument is a text constant, stated or stored. This subcommand must be given twice.

BY C...C

Standardizes the combination of two worksheets. The order of the data within the BY columns becomes ascending. The length and order of the remaining columns in the merged worksheet are manipulated according to the BY column adjustments, no matter the order or length of the original columns.

The length of the columns also depends on how you would like to handle multiple, unmatched, and missing observations within the BY columns. The adjusted length of the BY column or columns is still applied to the order and length of the remaining columns.

Requirements for BY columns are as follows:

- The column names for each of the two worksheets do not have to be unique. Minitab will create a unique name for the columns in the merged worksheet based on the original worksheet names.
- You must specify at least one pair of BY columns (one column from each worksheet) in the BY Columns sub-dialog box. Additional BY columns must also come in pairs.
- BY columns must have the same data type (numeric, date/time, or text).
- When merging columns with value ordered text, the value orders in the BY columns must be the same for both worksheets. For more information on value ordering, go to VORDER: Session command for controlling the order for text categories to be processed by Minitab commands on page 138.
- Column lengths must be the same between multiple BY columns.
- Columns that are not the same lengths as the BY columns are excluded from the merged worksheet. A note in the output documents the number of columns excluded.

NOMULTIPLES

Ignores all but the first row with the same values for the BY columns. If more than one row in a worksheet has values for the BY columns, then only the first such row is used.

Including multiple observations indicates that you wish to keep the observations of the BY column that are repeated within the column, regardless of the BY column values in the other worksheet. However, the values can still be matched to multiple observations in the BY columns of the other worksheet, depending on whether or not NOMULTIPLES is used for the other worksheet. Therefore, you could end up with longer columns in the merged worksheet. For example, if there are two rows in one worksheet and three rows in the other, all with the same values of the BY columns, then six rows would result.

NOMULTIPLES is available for each of the worksheets to be merged.

NOUNMATCHED

Removes unmatched BY column rows from the merged worksheet. If any row in a worksheet has values for the BY columns that are not matched in the BY columns for the other worksheet, then that row is not used.

By default, unmatched values are included which means that all of the rows in the opposite worksheet that do not have matching BY column values are included. Missing value symbols would be added to the cells for the entire row associated with the unmatched BY column values for the opposite worksheet. Consequently, padding rows with missing values will make all of the columns for both of the worksheets the same length as the BY columns.

NOUNMATCHED is available for each of the worksheets to be merged.

INCLUDE C...C

Specifies which columns to include from the worksheet, whether you merge worksheets using the default settings or using BY Columns. By default, all columns are included.

MISSINGS

Includes missing values within BY columns for both worksheets. The missing observations will be treated as distinct values. The missing BY column value will be matched with a missing BY column value in the other worksheet.

Note Missing values in text columns are represented by blanks and missing values in numeric and date/time columns are represented by an asterisk.

SORT: Session command for sorting columns

SORT [C...C C...C]

Sorts one or more columns.

The default is to sort by the first column and carry along additional columns. Sorting by multiple columns is done with the BY subcommand. SORT handles any combination of alpha or numeric columns.

The following is a simple example of the default method. Sorting is done based on the first column specified, C2. The next two columns, C3 and C4, are carried along. Sorting is done in ascending order unless you use the subcommand DESCENDING.

```
SORT C2 C3 C4 C12 C13 C14
```

Before sorting

C2	C3	C4
2	10	-1
3	11	-1
1	12	-3
4	13	-1
5	14	-1

After sorting

C12	C13	C14
1	12	-3
2	10	-1
3	11	-1
4	13	-1
5	14	-1

BY C...C

Specifies the columns to use to sort the worksheet. Rows are first sorted by the first column listed following BY, then, within that, by the second column, then, within that, by the third, and so on.

```
SORT C2 C3 C4 C12 C13 C14;
  BY C2.
```

Before sorting with BY

C2	C3	C4
2	10	-1
3	11	-1
1	12	-3
4	13	-1
5	14	-1

After sorting with BY

C12	C13	C14
1	12	-3
2	10	-1
3	11	-1
4	13	-1
5	14	-1

DESCENDING C...C

Requests that sorting be done in descending, rather than the default ascending, order. Columns listed on DESCENDING must also be listed on BY, or must be the first column listed on SORT if no BY is used.

```
SORT C3 C23;
  BY C2;
  DESCENDING C2.
```

Before sorting with DESCENDING

C2	C3
2	10
3	11
1	12
4	13
5	14

After sorting with DESCENDING

C23
14

C23
13
11
10
12

UNEQUAL

If the columns to be sorted have different numbers of rows, adds missing values so that the columns have an equal number of rows.

Storage

Important You cannot use NEWWS and AFTER local macros. For more information, go to Session commands that are not allowed in macros on page 1262.

Note These subcommands are optional, and are mutually exclusive. If you do not use one of these subcommands, results are stored in the columns that you specify with the main command.

NEWWS

Stores the results in a new worksheet, with the default name.

NEWWS *["text"]*

NEWWS [K]

Stores the results in a new worksheet. If you do not specify a worksheet name in a "text" argument or in a constant (K) argument, then Minitab uses the default naming of worksheets.

AFTER *["text"]*

Stores the results at the end of the specified worksheet. If you do not specify a worksheet name in a "text" argument or in a constant (K) argument, then Minitab stores the copied data at the end of the active worksheet.

ORIGINAL

Stores the results in the original columns.

RANK: Session command for ranking values in a column

RANK

Calculates and stores the ranks of the input column. Assigns the numeral 1 to the smallest value, the numeral 2 to the second smallest value, the numeral 3 to the third smallest value, and so on. Ties are assigned the average rank.

Note RANK works only with numeric columns.

The following command language ranks the values in C1 and puts the ranked values in C2.

```
RANK C1 C2
```

Before ranking

C1
0.5
1.0
1.5
1.0
2.0
0.0

After ranking

C1	C2
0.5	2.0
1.0	3.5
1.5	5.0
1.0	3.5
2.0	6.0
0.0	1.0

DELETE: Session command for deleting rows of data

DELETE K...K C...C

Deletes rows K...K from columns C...C, and moves the remaining rows up to close the gap. DELETE works with both text and numeric columns.

For example, the following command language changes the worksheet as shown below.

```
DELETE 2 5 6 C2-C4
```

Before DELETE

C2	C3	C4
23	4	154
31	5	175
22	4	143
26	3	153
32	6	167
30	6	158
24	4	155

After DELETE

C2	C3	C4
23	4	154
22	4	143
26	3	153
24	4	155

You can abbreviate a list of consecutive rows by using a colon. For example, to delete rows 1 through 10 and rows 25 through 30 from C1, use the following command language.

```
DELETE 1:10 25:30 C1
```

SPLIT: Session command for splitting a worksheet into multiple worksheets

SPLIT

Important SPLIT cannot be used in local macros. For more information, go to Session commands that are not allowed in macros on page 1262.

The SPLIT command splits a worksheet into multiple new worksheets, one for each combination of a set of BY variables. Each new worksheet is automatically named to reflect the combination of the BY variables.

Use SUBSET on page 115 to copy specified rows from the active worksheet to a new worksheet. With this command, you can specify the subset based on row numbers, brushed points on a graph, or a condition such as unmarried males under 50 years old.

BY C...C (required)

Specifies the columns to use to subset the worksheet. A new worksheet will be created for each unique value or combination of values in the columns. Columns must be non-empty, all the same length, and can be of any data type.

COLUMNS C...C

Specifies the columns to be copied to the new worksheet. Columns must be the same length as the BY columns. Columns other than these are not copied.

MATRICES *M...M*

Specifies the matrices to be copied to the new worksheet. Matrices must has the same number of rows as the BY columns. Matrices other than these are not copied.

NOMATRICES

Prevents matrices from being copied to the new worksheets.

NOCONSTANTS

Prevents stored constants from being copied to the new worksheets.

MISSINGS

Specifies that missing values are treated as a distinct value of the BY variable.

DATE: Session command for changing data type to date/time

DATE C...C C...C

Changes data type of C...C to date/time and places results in C...C.

FORMAT

Indicates the format of the new date/time column or columns. For example:

- `FORMAT (DTm/d/yy)` formats a new column as 5/25/16
- `FORMAT (DTm-d-yy)` formats a new column as 5-25-16

For multiple columns, add the number of columns to the format statement. For example, for 2 columns, use FORMAT (2DTm-d-yy).

For more information on date/time formats, go to Default date/time formats on page 1223.

ERASE: Session command for erasing variables

ERASE *E...E*

Erases any combination of columns (including their names), constants, and matrices.

Erasing all variables that you no longer need is a good practice.

ROWTOC: Session command for stacking multiple columns into one column

ROWTOC C...C C

Stacks rows in C...C and places them in C. Enter the columns that contain the data you want to appear in a single column, and then enter a column number or name in which to store the stacked data.

EXPAND C...C

Expands the specified column or columns, while stacking the rows.

EXSTORE C...C

Stores the expanded columns in the specified column.

SUBSCRIPTS C

Stores row subscripts in the specified column.

CSUBS C

Stores column subscripts in the specified column.

SUBSET: Session command for copying specified rows to a new worksheet

SUBSET

Important You cannot use SUBSET in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Use SUBSET to copy specified rows from the active worksheet to a new worksheet. With this command, you can specify the subset based on rows that match conditions, row numbers, formatted rows, brushed points on a graph, or a formula that you specify, such as unmarried males under 50 years old.

One of INCLUDE and EXCLUDE must be used.

One of BRUSHED, WHERE, ROWS, conditions for values in a column, or conditions for row numbers, must be used. If more than one is used, only the last valid one is honored.

Use SPLIT on page 113 to split, or unstack, the active worksheet into two or more new worksheets based on one or more "By" variables. SUBSET and SPLIT always copy data to new worksheets.

INCLUDE

Includes the specified rows.

INCLUDE and EXCLUDE are mutually exclusive: the last one issued is honored. Choose INCLUDE or EXCLUDE based on the degree of subsetting. For example, if you wish to use all but a few select rows in creating your graph, it is more efficient to use EXCLUDE and name a small number of rows with the ROWS subsubcommand.

EXCLUDE

Excludes the specified rows.

INCLUDE and EXCLUDE are mutually exclusive: the last one issued is honored. Choose INCLUDE or EXCLUDE based on the degree of subsetting. For example, if you wish to use all but a few select rows in creating your graph, it is more efficient to use EXCLUDE and name a small number of rows with the ROWS subsubcommand.

MINCLUDE

Includes rows that contain missing values in the subset column.

MINCLUDE and MEXCLUDE are mutually exclusive: the last one issued is honored.

MEXCLUDE

Excludes rows that contain missing values in the subset column.

MINCLUDE and MEXCLUDE are mutually exclusive: the last one issued is honored.

AND

Allows you to specify multiple conditions. For example, LT C1 50 AND EQUAL C2 'Blue' denotes rows that contain 50 or less in C1 and contain 'Blue' in C2.

OR

Allows you to specify multiple conditions.

Conditions for values in a column

These subcommands specify values in a column. For example:

- `ANY C1 22 23 27` specifies the rows in C1 that contain 22, 23, and 27.
- `NEQUALS C1 'male'` specifies the rows in C! that do not contain 'male'.

ANY C K...K
Values in C that equal any of the specified values.

EQUALS C K
Values in C that equal the specified value.

NEQUALS C K
Values in C that do not equal the specified value.

Conditions for values in a numeric or date/time column

These subcommands specify values in a numeric or date/time column. For example:

- `GT C1 30` specifies the values in C1 that are greater than 30.
- `LTLT C1 30 70` specifies the values in C1 that are greater than 30 and less than 70.

LT C K
Values in C that are less than K.

LE C K
Values in C that are less than or equal to K.

GT C K
Values in C that are greater than K.

GE C K
Values in C that are greater than or equal to K.

LELE K C K
Values in C that are greater than or equal to K and less than or equal to K.

LELT K C K
Values in C that are greater than or equal to K and less than K.

LTLE K C K
Values in C that are greater than K and less than or equal to K.

LTLT K C K
Values in C that are greater than K and less than K.

Conditions for values in a text column

These commands specify values in a text column. The text value must be in single quotation marks. For example:

- `BEGINS C1 'm'` specifies all values in C1 that begin with "m".
- `NCONTAINS C1 'non'` specifies all values in C1 that contain "non".

BEGINS C K
Values in C that begin with K (text columns only).

ENDS C K
Values in C that end with K (text columns only).

CONTAINS C K
Values in C that contain K (text columns only).

NCONTAINS C K
Values in C that do not contain K (text columns only).

Conditions for row numbers

ROWS K...K
Specifies row numbers K through K. List all row numbers to be included or excluded with a space between each. Denote a patterned range by K:K/K. For example, `10:50/5` denotes all values from 10 to 50 by intervals of 5.

ROWLE K
The row number is less than or equal to K.

ROWGE K
The row number is greater than or equal to K. For example, `ROWGE 12` denotes row numbers that are greater than or equal to 12.

ROWBETWEEN K K
The row number is between K and K inclusive. The first value must be less than or equal to the second.

Other conditions

BRUSHED
Specifies the rows that are currently brushed.

WHERE *expression*
Specify a formula for subsetting. This formula must be in quotes. For example:

- To subset where C1 is 2, enter `WHERE "C1 = 2"`.
- To subset where C2 = a, enter `WHERE "C2 = ""a"""`.
- To subset where C3 = 3/3/16, enter `WHERE "C3 = DATE(""3/3/98"")"`.

CFORMAT C
Specifies a column that includes formatted rows.

Options for the new worksheet

NAME *"text"*

NAME K
Specifies the name of the new worksheet in a "text" argument or in a constant (K).

COLUMNS C...C
Specifies one or more columns of the appropriate length to be copied to the new worksheet. Columns other than these are not copied.

EQLN
Includes in the subset only columns that are the same length as the column with the condition. If this subcommand is not used, all columns in the worksheet are included in the subset.

MATRICES *M...M*
Specifies one or more matrices with the appropriate number of rows to be copied to the new worksheet. Matrices other than these are not copied.

NOMATRICES
Prevents matrices from being copied to the new worksheet.

NOCONSTANTS
Prevents stored constants from being copied to the new worksheet.

READ data into columns

READ C...C
Reads in data, row by row, that you type from the keyboard, or that you import from a text file. You cannot type comments on a data line when the FORMAT subcommand is used.

READ enters new data into columns, replacing any data already in those columns, if it exists. For information on entering data into a matrix, go to READ data into a matrix on page 44.

When you enter data manually, type `END.` after you enter your final value.

When you use READ, you can use a space or a comma to separate data entries. For example:

```
READ C1 C5.
1 2
3,4
END.
```

For details on using this command without subcommands to select data entry options, go to Using READ without subcommands on page 1266.

FILE *"filename"*
Inserts data from the specified text file. You may specify the filename as either the name of the file in double quotes, or a stored text constant. If the file has an extension other than DAT and/or if it is not in your current directory, include the file extension and the path within the single quotation marks. For example, use the following command to read a copy of the file SALES.ASC stored in the subdirectory JANUARY underneath the directory SMITH on the C drive.

```
READ C1-C5;
 FILE "C:\SMITH\JANUARY\SALES.ASC".
```

FORMAT *(format statement)*
Include a format statement, within parentheses to specify precisely how to enter data into the worksheet. The entire expression within the parentheses is repeated once for each record. For more information, go to Using READ with FORMAT on page 1266.

The FORMAT subcommand is useful when you want to skip over spaces, read data that have no spaces between them, insert decimal points in numbers, or read in text data, date/time data, or currency data.

Format items may be combined together. For example, the following command reads the name in the first 20 spaces of each data line into Name (C12), skips the next 10 spaces (spaces 21 through 30), then reads the number in space 31 into C1, the number in space 32 into C2, ..., the number in space 40 into C10.

```
NAME C12 'Name'
READ 'Name' C1-C10;
```

```
 FILE "MYDATA";
 FORMAT(A20, 10X, 10F1).
```

Minitab has a special date/time (DT) format which works as shown below. This says to read the date/time value in the first 8 spaces in the file into C1, and that the format of the date/time data in the file is m/d/yy.

```
READ C1;
 FILE "DATEDATA";
 FORMAT(DT8m/d/yy).
```

The following example shows the use of a decimal indicator, repeat factor in front of parentheses, and the slash.

```
READ C11-C15;
 FILE "EMPLOYEEDATA";
 FORMAT (F2.1, 2(1X,F3), F4/F2).
```

This example uses two data lines for every row read. From the first line, the value of C11 is in spaces 1 and 2. The first value is a whole number and the second value is in the tenths place. The format skips space 3. Then, C12 is read from spaces 4 to 6. The format skips space 7. Then, C13 is read from spaces 8 to 10, which repeats the pattern inside of the parentheses. C14 is read from spaces 11 to 14. In response to the /, reading moves to the second data line, and C15 is read from spaces 1 and 2. For more information, go to Valid format items on page 1269.

NOBS K

The NOBS subcommand specifies the number of observations (rows) to be inserted. If an END on page 39 subcommand or end-of-file is encountered before K observations are inserted, NOBS is ignored. NOBS is useful when you want to insert just the first portion of a file. It is also useful for Prompting a user for information on page 1260.

SKIP K

Tells Minitab to skip K lines at the top of the data file before beginning to add data into the file. This is most useful when you have one or more lines of text, such as column names and titles, at the top of a data file that you want to import into Minitab.

With READ; FILE only

DECIMAL works with READ when only when reading a file. DECIMAL does not work with READ if you type data.

DECIMAL ","

DECIMAL "."

Specifies a comma or period as a decimal separator.

READ data into a matrix

READ K K M

Puts numbers into a matrix. To input data to columns, go to READ data into columns on page 42. You can specify the filename as either the name of the file in double quotes, or a stored text constant. If the file has an extension other than DAT and/or if it is not in your current directory, include the file name extension and the path within quotation marks.

You can use either spaces or commas to separate the data in the matrix.

You must specify the dimension of the matrix in the READ command. The first K gives the number of rows, the second K the number of columns. The M is the matrix identifier for storage. If a file name is not used, READ is

followed by data lines, each containing one row of the matrix. The following command creates the following matrix.

Command

```
READ  3  4  M2
1  2  3  4
5  6  7  8
9  10  11  12
END
```

Matrix

```
1  2  3  4
5  6  7  8
9  10  11  12
```

FILE *"filename"*

FILE K

Reads or inserts data from the specified text file.

CONCATENATE: Session command for combining text columns

CONCATENATE C...C C

Combines text columns C...C and places into C, to form longer words.

In the following example, portions of numbers were originally entered into columns C1–C3. CONCATENATE combines C1-C3 into one column, C4.

```
READ C1-C3;
 FORMAT (A4, A3, A4).
192-42-7777
123-45-6789
END
CONCATENATE C1-C3 C4
```

The worksheet is changed as follows.

C1-T	C2-T	C3-T	C4-T
192-	42-	7777	192-42-7777
123-	45-	6789	123-45-6789

CODE: Session command for changing values in columns to new values

CODE (K...K K ... (K...K)K C...C [C...C]

Copies columns, changing the indicated values. For example, the following command language codes every –1 and –2 to 100.

```
CODE (-1 -2)100 C1 C2;
 AFTER.
```

The command language places the results at the end of the worksheet (C2), as follows.

C1	C2
-1.0	100.0

C1	C2
3.0	3.0
1.1	1.1
-1.0	100.0
0.0	0.0
-2.0	100.0
2.4	2.4

An interval can be abbreviated with a colon. For example, the following command language changes all values in the range 1 to 1.5, and the value 2 into a 5.

```
CODE (1:1.5, 2)5 C1-C3 C11-C15
```

You can make several changes at one time. For example, suppose the data in columns C1–C5 include integers 1 through 10. In this case, the following command language changes all values in C1–C5 that are from 1 to 5 into 10, and all values that are from 6 to 10 into 20, and stores the results in C11–C15.

```
CODE (1:5)10 (6:10)20 C1-C5 C11-C15
```

CODE can change values into the missing data code. For example, to change –99 to *, use the following command language.

```
CODE (-99) '*' C1-C10;
ORIGINAL.
```

You can also code text data. Enclose text values with double quotation marks. Denote a missing text value as two double quotation marks, with no space in between, as "". For example, use the following command language to code 1, 2, and 3 to low, medium, and high.

```
CODE (1) "low" (2) "medium" (3) "high" C1 C2
```

You can also use CONVERT on page 122 to recode text data.

ENDPOINTS K

For coding ranges of values, the value of K controls inclusion of the endpoints.

Value of K	Endpoints that are included in the range
1	Lower endpoint only
2	Upper endpoint only
3 (default)	Both endpoints
4	Neither endpoint

TSUMMARY

Shows the summary table in the output.

TSUMMARY and NODEFAULT are mutually exclusive.

NODEFAULT

Suppresses the summary table in the output.

TSUMMARY and NODEFAULT are mutually exclusive.

LFPERCENT K K
Recodes the unique values (of the input columns specified in the CODE command) above the K^{th} percentile in count to K, and places the recoded values into output columns specified in the CODE command.

LFVALUE K K
Recodes the unique values (of the input columns specified in the CODE command) that occur less than K times to K, and places the recoded values into output columns specified in the CODE command.

Storage

The following subcommands are optional, and are mutually exclusive. If you do not use one of these subcommands, results are stored in the columns that you specify with the main command.

Note You cannot use NEWWS and AFTER in a local macro.

NEWWS

NEWWS *"text"*

NEWWS K
Stores the results in a new worksheet. If you do not specify a worksheet name in a "text" argument or in a constant (K) argument, then Minitab uses the default naming of worksheets.

AFTER

AFTER *"text"*
Stores the results at the end of the specified worksheet. If you do not specify a worksheet name in a "text" argument, then Minitab stores the copied data at the end of the active worksheet.

ORIGINAL
Stores the results in the original columns.

CONVERT: Session command for converting text data to numeric data, and numeric data to text data

CONVERT C C C C
Using the conversion table in C C, converts C to C.

A conversion table, assigning a numeric value to each text value, must be put into two columns before CONVERT is used. Then, using these matching values, CONVERT changes all corresponding numeric values to the correct text values (or vice versa). If no match is found, a missing value is stored. Missing values are denoted by * in a numeric column and a blank in a text column. CONVERT can also be used to convert from numeric to numeric or from text to text.

Use the following commands to store a conversion table in C1–C2. For more information on entering data using format statements, go to READ data into a matrix on page 44 or READ data into columns on page 42.

```
READ C1 C2;
  FORMAT (F1, 1X, A6).
1 RED
2 YELLOW
END
```

Use the conversion table to do two conversions. The first CONVERT converts numbers in C3 to colors in C10. The second converts colors in C4 to numbers in C11.

```
CONVERT C1 C2  C3 C10
CONVERT C2 C1  C4 C11
```

The worksheet is changed as follows.

C1	C2-T	C3	C4-T	C10-T	C11
1	RED	1	RED	RED	1
2	YELLOW	2	RED	YELLOW	1
		1	YELLOW	RED	2
		3	RED		1
		2	BLUE	YELLOW	*

NODEFAULT

Suppresses the summary table in the output.

DROUND: Session command for rounding date/time values

DROUND C C

Rounds the original date/time values in the first C and stores them in the second C.

For example, suppose you have sales data for each day of the year, but you want to summarize sales for each quarter of the year. You can use Round Date/Time to round the dates down to the nearest quarter.

Units for rounding

Specify the unit to round to. You must use one of these mutually exclusive subcommands.

YEAR

QUARTER

MONTH

WEEK

DAY

HOUR

MINUTE

SECOND

TENTH

HUNDREDTH

THOUSANDTH

Options

DROUND rounds down to the nearest whole unit.

UP
Use UP to round up to the nearest whole unit.

NEAREST
Use NEAREST to round up or down to the nearest whole unit.

NUMERIC: Session command for changing the data format of a date/time column or extracting date/time components

NUMERIC C C
The NUMERIC command serves two functions:

- To change a date/time or a text column to a numeric column. (Use with a date/time or text input column.)
- To extract one or more components (such as the day, quarter, or hour) from a date/time column, and save those components in a numeric column. (Use with a date/time input column.)

Note When the input column is text, no subcommands are available.

Date/time input

WKDAY
Extracts the day of the week (Sun, Mon, Tue, Wed, Thu, Fri, Sat).

DAY
Extracts the day of the month (01, 02, ..., 31).

WEEK
Extracts the week of the year number (Wk01 - Wk53).

MONTH
Extracts the month (Jan, Feb, Mar, Apr, ..., Dec).

QUARTER
Extracts the quarter (Q1, Q2, Q3, Q4).

YEAR
Extracts the year.

TWODIGIT
Uses the two digit format (00, 01, ..., 99).

FOURDIGIT
Uses the four digit format (2000, 2001, ..., 2099).

HOUR
Extracts the hour (00, 01, ..., 23).

MINUTE
Extracts the minute (00, 01, ..., 59).

SECOND
Extracts the second (00, 01, ..., 59).

TENTHS
Extracts the tenths of a second (0, 1, 2, ..., 9).

HUNDREDTHS
Extracts the hundredths of a second (0, 1, ..., 99).

THOUSANDTHS
Extracts the thousandths of a second (0, 1, 2, ..., 999).

TEXT: Session command for changing the data type of a column to text

TEXT C...C C...C
Changes data type of C...C to text and places results in C...C. The TEXT command serves the following functions:

- Use TEXT to change a date/time or a numeric column to a text column. (Use with a date/time or numeric input column.)
- Use TEXT to extract one or more components (such as the day, quarter, or hour) from a date/time column, and save those components in a text column. (Use with a date/time input column.)
- Use TEXT to format the width of a text column. (Use with a text input column.)

Minitab automatically formats the new alpha column up to 8 characters wide. Use WIDTH or MAXWIDTH (you cannot use both) to override this default.

Date/time input column

WKDAY
Extracts the day of the week (Sun, Mon, Tue, Wed, Thu, Fri, Sat).

DAY
Extracts the day of the month (01, 02, ..., 31).

WEEK
Extracts the week of the year number (Wk01 - Wk53).

MONTH
Extracts the month (Jan, Feb, Mar, Apr, ..., Dec).

QUARTER
Extracts the quarter (Q1, Q2, Q3, Q4).

YEAR
Extracts the year.

TWODIGIT
Uses the two digit format (00, 01, ..., 99).

FOURDIGIT
Uses the four digit format (2000, 2001, ..., 2099).

HOUR
Extracts the hour (00, 01, ..., 23).

MINUTE
Extracts the minute (00, 01, ..., 59).

SECOND
Extracts the second (00, 01, ..., 59).

TENTHS
Extracts the tenths of a second (0, 1, 2, ..., 9).

HUNDREDTHS
Extracts the hundredths of a second (00, 01, ..., 99).

THOUSANDTHS
Extracts the thousandths of a second (000, 001, 002, ..., 999).

WIDTH K
Fixes the width (in characters) of the output column.

MAXWIDTH K
Fixes the maximum width (in characters) of the output column.

MISSING *"text"*
Converts missing values to "text".

Numeric input column

WIDTH K
Fixes the width (in characters) of the output column.

MAXWIDTH K
Fixes the maximum width (in characters) of the output column.

SIGNIFICANT K
Specifies the number of significant digits to maintain (up to a maximum of 6). You cannot use both SIGNIFICANT and DECIMALS with one command.

DECIMALS K
Specifies the number of decimal digits to maintain. You cannot use both SIGNIFICANT and DECIMALS with one command.

MISSING *"text"*
Converts the missing value symbol * to a specified character string. Unless you specify the width of your output with WIDTH or MAXWIDTH, the default length for the string is 8 characters.

Text input column

WIDTH K
Fixes the width (in characters) of the output column.

MAXWIDTH K
Fixes the maximum width (in characters) of the output column.

STACK: Session command for stacking blocks of columns and constants on top of each other

STACK *(E...E)* C

STACK *(E...E)* ... *(E...E)* (C...C)
Stacks blocks of columns and constants on top of each other.

The following command language stacks C3 and C4 on C1 and C2.

```
STACK (C3 C4) (C1 C2) (C5 C6)
```

Columns to stack

C1	C2	C3	C4
66	130	70	155
64	125	66	145
65	115	69	160

Stacked columns

C5	C6
70	155
66	145
69	160
66	130
64	125
65	115

The following command language puts 66, 64, 65, 70, 66, 69, 67, and 63 into C10. The last column (or block of columns) specified is the target column.

```
STACK C1 C3 67 63  C10
```

SUBSCRIPTS C
Creates a column of subscripts in the original worksheet. The first block is given the subscript 1, the second block the subscript 2, and so on.

USENAMES
Creates subscripts based on variable names.

NEWWS

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in a "text" argument or in a constant (K) argument, then Minitab uses the default naming of worksheets. Minitab automatically creates a subscripts column in the new worksheet.

Important You cannot use NEWWS in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

USENAMES

Creates subscripts based on variable names.

UNSTACK: Session command for separating a column into multiple columns

UNSTACK (C...C) *(E...E)...(E...E)*

Separates one or more columns into several blocks of columns and stored constants, based on subscript values stored in an associated column.

Unstack is useful for subsetting data.

SUBSCRIPTS C...C

Uses columns to unstack data. The rows with the smallest subscript are stored in the first block, the rows with the second smallest subscript in the second block, and so on. If you do not use SUBSCRIPTS, each row is stored in a separate block.

In general, each block must be enclosed in parentheses. However, if a block contains just one argument, you may omit the parentheses.

For most applications, the subcommand SUBSCRIPTS is needed.

MISSINGS

Includes rows subscripted with missing data as part of the unstacked data.

NEWWS

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in "text" or in a constant (K) argument, then Minitab uses the default naming of worksheets.

Important You cannot use NEWWS in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

VARNAMES

Labels the columns for the unstacked data with the source variable name (the source column's label or number if no label exists) and the corresponding row entries from the subscript column or columns. If you do not use SUBSCRIPT, the text strings "Row1", "Row2", and so on are used.

AFTER

Appends the unstacked data into empty columns after the last column of data in the current worksheet.

Important You cannot use AFTER in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

WSTACK: Session command for stacking worksheets

WSTACK

Important You cannot use WSTACK in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

Use WSTACK to stack two or more worksheets. The argument is the names of the worksheets. For example, WSTACK "Worksheet1" "Worksheet 2" stacks Worksheet 1 on top of Worksheet 2. Minitab stacks columns with the same names. Columns that don't have the matching names are kept as separate columns.

Within this command, you can specify that the results are placed in a new worksheet (NEWWS) or appended to an existing worksheet (APPEND). NEWWS and APPEND subcommands are mutually exclusive. If neither is specified, NEWWS is assumed.

You can use MERGE on page 107 to combine worksheets side-by-side.

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in "text" or in a constant (K) argument, then Minitab uses the default naming of worksheets.

APPEND

Appends the data to the bottom of the first worksheet that you specified in WSTACK.

SOURCE *["column name"]*

SOURCE appends the source worksheet name to the bottom of the specified column. If you do not specify the column, the command creates a new ID column in the output sheet.

TSUMMARY (default)

TSUMMARY and NODEFAULT are mutually exclusive. TSUMMARY shows the summary table in the output.

NODEFAULT

TSUMMARY and NODEFAULT are mutually exclusive. NODEFAULT suppresses the summary table in the output.

TRANSPOSE: Session command for changing rows to columns, and columns to rows

TRANSPOSE C...C

TRANSPOSE C *M*

TRANSPOSE *M* C

TRANSPOSE *M [M]*

TRANSPOSE reconfigures data so that rows become columns and columns become rows.

- TRANSPOSE C...C transposes the values in the columns.
- TRANSPOSE C M transposes the values in a column to a matrix.
- TRANSPOSE MC transposes the values in a single row matrix to a column.
- TRANSPOSE M [M] transposes the values in a matrix to another matrix.

VARNAMES C

Specifies a column that contains variable names for the transposed columns.

STORE C...C

STORE *M*

Specifies where the transposed data are to be written. You can store the data in columns or a matrix. When you transpose data to columns, the number of columns specified must equal the number of rows in the original columns.

If you do not use STORE, the data are stored in a new worksheet.

LABELS C

Specifies the column in which the labels of the transposed column or columns or matrix are stored.

NEWWS

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in "text" or in a constant (K) argument, then Minitab uses the default naming of worksheets.

Important You cannot use NEWWS in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

AFTER

Appends the transposed data into empty columns after the last column of data in the current worksheet.

Important You cannot use AFTER in a local macro. For more information, go to Session commands that are not allowed in macros on page 1262.

COPY: Session command for copying data

COPY ***[E...E E...E]***

Copy specified columns, constants, and matrices, or a subset thereof, into a similar or compatible format in a specified worksheet.

In addition to format, the source and destination must be compatible in size. For example, you cannot copy a column with 100 rows into a set of 50 constants. You can subset the source data to accommodate the size of the destination.

If you don't specify an argument, COPY copies the entire contents of the current worksheet to another worksheet. You can specify an argument to do the following:

- Copy columns into columns, constants, or a matrix.
- Copy constants into constants or columns.
- Copy a matrix into columns or a matrix.

Specify a subset

INCLUDE and EXCLUDE are mutually exclusive. The last one issued is honored. Choose INCLUDE or EXCLUDE based on the degree of subsetting. For example, if you want to use all rows, except a few, use EXCLUDE and specify a small number of rows with the ROWS subsubcommand.

INCLUDE

Includes the specified rows.

BRUSHED

Use to specify all brushed rows for subsetting.

WHERE ***"expression"***

Specify a condition to be met for subsetting. This condition must be in double quotation marks. For example:

- To subset where C1 is 2, enter `WHERE "C1 = 2"`.
- To subset where C2 = a, enter `WHERE "C2 = ""a""`.
- To subset where C3 = 3/3/16, enter `WHERE "C3 = DATE(""3/3/16"")"`.

ROWS K:K

List all row numbers to be included or excluded with a space between each. Denote a patterned range by K:K/K. For example, 10:50/5 denotes all values from 10 to 50 by intervals of 5.

EXCLUDE

Excludes the specified rows.

BRUSHED

Use to specify all brushed rows for subsetting.

WHERE ***"expression"***

Specify a condition to be met for subsetting. This condition must be in double quotation marks. For example:

- To subset where C1 is 2, enter `WHERE "C1 = 2"`.
- To subset where C2 = a, enter `WHERE "C2 = ""a""`.
- To subset where C3 = 3/3/16, enter `WHERE "C3 = DATE(""3/3/16"")"`.

ROWS K:K

List all row numbers to be included or excluded with a space between each. Denote a patterned range by K:K/K. For example, 10:50/5 denotes all values from 10 to 50 by intervals of 5.

Specify target location

Important NEWWS, AFTER, and STORE cannot be used in local macros. For more information, go to Session commands that are not allowed in macros on page 1262.

Note You can use only one NEWWS, AFTER, or STORE subcommand.

NEWWS

NEWWS *"text"*

NEWWS K

Stores the results in a new worksheet. If you do not specify a worksheet name in "text" or in a constant (K) argument, then Minitab uses the default naming of worksheets.

AFTER

AFTER *"text"*

AFTER K

Stores the results at the end of the specified worksheet. If you do not specify a worksheet name in a "text" argument or in a constant (K) argument, then Minitab stores the copied data at the end of the active worksheet.

STORE K *E...E*

STORE *"text" E...E*

Stores the copied data in the specified columns constants, or matrices of the worksheet named in a "text" argument or in a constant (K).

Naming

VARNAMES

Gives default names to the stored columns, constants, or matrix.

RMERGE: Session command for merging worksheets

RMERGE K...K

Use RMERGE to combine two or more worksheets side-by-side into a new worksheet. The argument is the names of the worksheets. For example, `RMERGE "Worksheet1" "Worksheet 2"` merges Worksheet 2 to the right of Worksheet 1.

Note Stored constants, matrices, DOE objects, and worksheet descriptions do not transfer into the merged worksheet.

NAME K (optional)

Specifies the name K for the new worksheet.

Editor

CFORMAT: Session command for conditional formatting of worksheet cells

CFORMAT C...C

Formats cells in the specified columns based on the specified rule.

You can specify only one rule subcommand (and possibly its sub-subcommands) with the CFORMAT command at a time. However, you can run CFORMAT on the same column multiple times, and the rules accumulate. If a cell is affected by multiple rules, the formatting of the most recent rule is used.

The availability of a specific rule might depend on the data type of the column specified in CFORMAT.

Attribute subcommands

The following subcommands define the cell format for cells that meet the condition.

TCOLOR K

Specifies the text color with a number value in K, which is a number from Numbers for colors to use in session commands on page 1255.

BOLD

Applies bold style to text in the cells to format.

NOBOLD

Removes bold style from text in the cells to formats.

ITALIC

Applies italic style to text in the cells to format.

NOITALIC

Removes italic style from text in the cells to format.

UNDERLINE

Underlines text in the cells to format.

NOUNDERLINE

Removes underline from text in the cells to format.

COLOR K

Specifies the fill color of the cells with a number value in K, which is a number from Numbers for colors to use in session commands on page 1255.

Rules for highlighting cells

MISSING

Formats cells that contain missing observations (available for columns with any data type).

GT K

Formats cells that are greater than K (available for columns with any data type).

LT K
Formats cells that are less than K (available for columns with any data type).

BETWEEN K K
Formats cells between K and K, inclusive (available for columns with any data type).

NOTBETWEEN K K
Formats cells that are not between K and K, inclusive (available for columns with any data type).

EQUALS K
Formats cells that are equal to K (available for columns with any data type).

TEXT *"text"*
Formats cells that contain the text string, 'text" (available for text columns) The text string may be found within a larger text string. For example, if the rule is to format cells that contain "abc", a cell that contains dabcd would be formatted.

VALUES C

VALUES K...K
Formats cells that match any of the values in C or in K...K (available for columns with any data type) C or K...K have the same data type as the columns specified in CFORMAT.

NOVALUES C

NOVALUES K...K
Formats cells that do not match any of the values in C or in K...K (available for columns with any data type) C or K...K have the same data type as the columns specified in CFORMAT.

YESTERDAY
Formats cells that contain yesterday's date (available for date/time columns).

TODAY
Formats cells that contain today's date (available for date/time columns).

LAST K
Formats cells that contain a date that occurred within the last K days (available for date/time columns).

LWEEK
Formats cells that contain a date that occurred last week (available for date/time columns).

TWEEK
Formats cells that occur this week (available for date/time columns).

LMONTH
Formats cells that occurred last month (available for date/time columns).

TMONTH
Formats cells that occur this month (available for date/time columns).

Hi low rules

HVALUE K
Formats cells that contain the highest K values in the column (available for numeric and date/time columns). If there are ties at the K^{th} location, all tied values will be formatted.

HPERCENT K
Formats cells that contain the highest K percentage of values in the column (available for numeric and date/time columns).

LVALUE K
Formats cells that contain the lowest K values in the column (available for numeric and date/time columns) If there are ties at the K^{th} location, all tied values will be formatted.

LPERCENT K
Formats cells that contain the lowest K percentage of values in the column (available for numeric and date/time columns).

Pareto rules

MFREQUENT K
Formats cells that contain the K most frequent values (available for columns with any data type).

MPERCENT K
Formats cells that contain the top K% of values (available for columns with any data type).

BFREQUENT
Formats cells that contain the K least frequent values (available for columns with any data type).

BPERCENT K
Formats cells that contain the bottom K% of values (available for columns with any data type).

Statistical rules

OUTLIER
Formats cells that are outliers (available for columns with any data type) Columns can be different data types.

Outliers are determined based on one of the following sub-subcommands:

BOXPLOT
Observations that meet the conditions for outliers on a boxplot are formatted.

STDEV K
Observations that are more than K standard deviations away from the mean are formatted.

CONTROL K K
Formats cells that contain values that are out of control (available for numeric and date/time data).

The first K specifies the rule:

- K = 1 if data are continuous.
- K = 2 if the data are attribute.

The second K specifies the subgroup size. K > 0.

Examples of subcommand syntax:

- CONTROL 1 1 specifies to use I chart test 1.
- CONTROL 1 K, where K > 1, specifies to use XBar chart test 1.

- CONTROL K K, where K > 0, specifies to use P chart test 1.

The CONTROL subcommand honors preferences (**File** > **Options**) settings to determine the Test 1 rule criteria and to determine the estimate of standard deviation used.

SPECIFICATION

Formats cells that are outside of the specification limits (available for numeric and date/time data)

Specify one or both of the following:

LSPEC K

Specifies the lower specification limit. Cells with values less than K are out of specification.

K for LSPEC can be any numeric or date/time value, but K for LSPEC must be less that K for USPEC. Both Ks must match the type for the argument to CFORMAT.

USPEC K

Specifies the upper specification limit. Cells with values greater than K are out of specification.

K for USPEC can be any numeric or date/time value, but K for LSPEC must be less that K for USPEC. Both Ks must match the type for the argument to CFORMAT.

RESIDUAL

Formats cells that are unusual observations from a model because of large standardized residual (available for response columns that have a model associated with them).

UNUSUAL

Formats cells that are unusual observations from a model because they exhibit high leverage (available for response columns that have a model associated with them).

CFAUTOMATICALLY: Session command for automatically recalculating values

CFAUTOMATICALLY

Updates the data by automatically recalculating the values in columns or stored constants with assigned formulas whenever you change the corresponding data for variables. To update formulas manually, use CFMANUALLY on page 136.

This command cannot be used in a local macro.

CFMANUALLY: Session command for manually recalculating values

Allows you to manually recalculate the values in columns or stored constants with assigned formulas only when using the command CFNOW on page 137 (Calculate All Formulas Now). To update formulas automatically whenever the data they depend on changes, use CFAUTOMATICALLY on page 136.

This command cannot be used in a local macro.

CFNOW: Session command for recalculating values now

CFNOW

Recalculates the values in columns or stored constants with assigned formulas using the current data for variables. Can be used only when **Calculate All Formulas Automatically** is deselected. For more information about calculating formulas manually with command language, go to CFMANUALLY on page 136.

This command cannot be used in a local macro.

FDATE/TIME: Session command for changing the format of date/time columns

FDATE/TIME C...C

Changes the format of a date/time columns C...C. You can use FDATE/TIME for columns that are empty.

For more information on date/time formats, go to Default date/time formats on page 1223.

FORMAT *(format statement)*

Indicates the format of the date/time column. For example, use the subcommand FORMAT (DTm/d/yy) to format the column as 5/25/18. Use the subcommand FORMAT (DTm-d-yy) to format the new column as 5-25-18.

For multiple columns, add the number of columns to the format statement. For example, for 2 columns, use FORMAT (2DTm-d-yy).

FNUMERIC: Session command for changing columns to numeric format

FNUMERIC C...C

Changes the format of columns C...C to the specified numeric format. You can use FNUMERIC for columns that are empty.

AUTO

The values in the column determine the format of the column, for example, the number of decimal places and the currency symbol (if entered).

FIXED K

Displays numeric data to K decimal places.

EXPONENTIAL [K]

Displays numeric data using exponential notation. The optional argument K specifies the number of decimal places.

CURRENCY [K]

Displays currency data. The optional argument K specifies the number of decimal places, from 0 to 30 inclusive.

SYMBOL *"text"*

Specifies the currency symbol as a text value.

NEGATIVE K
Displays negative values using a minus sign (K = 1) or parentheses (K = 2).

PERCENTAGE K
Displays percentage data to K decimal places.

FTEXT: Session command for changing the format of text columns

FTEXT
Changes the format of text columns C...C. You can use FTEXT for empty columns.

FORMULA: Session command for assigning a formula to a column

FORMULA *E = expression*
Assigns a formula to a column or stored constant. The expression that defines the formula may contain arithmetic operations, comparison operations, logical operations, and functions. For a list of the functions you can use, go to LET: Session command for correcting a number in a worksheet or performing arithmetic on page 64.

This command cannot be used in a local macro.

RFORMULA: Session command for removing formulas

RFORMULA *E...E*
Removes formulas from the columns or constants that you specify. Removing a formula from a column does not remove the data from the column.

You cannot use this command in a local macro.

VORDER: Session command for controlling the order for text categories to be processed by Minitab commands

VORDERC...C
Use VORDER to control the order in which you would like text categories to be processed by Minitab commands. By default, text categories are processed in alphabetical order. However, alphabetical order might not always be the most convenient way to process your data.

VALUES K...K

VALUES C
Specifies the value order using either stored or stated text constants or a single text column.

If you use the VALUES subcommand more than one time, the last one with valid arguments wins. In the first form, stored or stated text constants are used to specify the values in order. In the second form, the rows of a single text column are used to specify the values in order. In either form, it is an error for values to be repeated. Note that the command language does not provide access to standard orderings.

WORKSHEET
Specifies to order data based on the first occurrence in the worksheet.

ALPHABETICAL
Specifies to order data in alphabetical order.

Basic Statistics

DESCRIBE: Session command for summarizing numeric data with statistics

DESCRIBE C...C

Produces descriptive statistics for each column. You can do the calculations and display graphs separately for each level of a BY variable.

For information on storing many different statistics, go to STATS on page 144.

Control of grouping

BY

Use BY to produce separate statistics and graphs for each unique value in C. The values in the BY column may contain numeric, text, or date/time data and may be any value. When the BY column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on. Column lengths must be equal to use BY.

Statistics

MEAN

Calculates the arithmetic mean, or average. The mean is a commonly used measure of the center of a batch of numbers.

Note Missing values are omitted from the calculation of the function Mean.

SEMEAN

Gives the standard error of the mean. It is calculated as StDev / SQRT (N).

STDEV

Calculates the sample standard deviation, which provides a measure of how spread out the data are. To calculate the variance, simply square the standard deviation value.

If the column contains x1, x2, ..., xn, with mean $\bar{x}$, then the standard deviation is calculated as follows:

$$\sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

Note Missing values are omitted from the calculation of the function Standard deviation.

VARIANCE

Variance is a measure of how far the data are spread about the mean. Sample variance equals the standard deviation squared. It can also be computed with this formula:

$$\frac{\Sigma(x-\bar{x})^2}{n-1}$$

CVARIATION

Displays the coefficient of variation, a measure of relative variability. It is calculated as: 100 (s / $\bar{x}$).

MEDIAN

Stores the median of a column. The median is in the middle of the data: half the observations are less than or equal to it. Suppose the column contains n values. If n is odd, the median is the value in the middle. If n is even, the median is the average of the two middle values.

Note Missing values are omitted from the calculation of the function Median.

MODE

Displays the mode of a column. The mode of a data set is the value that occurs most frequently. For example, the mode for data set {1, 3, 4, 4, 7} is 4.

Note For data sets that have multiple mode values, Minitab displays only the four smallest modes.

TRMEAN

A 5% trimmed mean is calculated. Minitab removes the smallest 5% and the largest 5% of the values (rounded to the nearest integer).

SUM

Calculates the sum.

Note Missing values are omitted from the calculation of the function Sum.

MIN

Stores the smallest number in a column.

Note Missing values are omitted from the calculation of the function MIN.

MAX

Stores the largest number in a column.

Note Missing values are omitted from the calculation of the function MAX.

RANGE

Calculates the difference between the largest and smallest data value.

SSQ

Squares each value in the column, and computes the sum of those squared values. That is, if the column contains x_1, x_2, ..., x_n, then sum of squares calculates ($x_1^2 + x_2^2 + ... + x_n^2$).

Note Missing values are omitted from the calculation of the function the function Sum of Squares (corrected).

SKEWNESS

Skewness is a measure of asymmetry. A value more than or less than zero indicates skewness in the data. But a zero value does not necessarily indicate symmetry.

KURTOSIS

Kurtosis is one measure of how different a distribution is from the normal distribution. A positive value characterizes a distribution with heavier tails than the normal distribution. A negative value characterizes a distribution with lighter tails than the normal distribution.

MSSD

MSSD computes half the Mean of the Squared Successive Differences of a batch of numbers. For example, suppose a column contains 1, 2, 4, and 10. The successive differences are 2 - 1 = 1, 4 - 2 = 2, and 10 - 4 = 6. Then:

MSSD = mean(1-squared, 2-squared, and 6-squared) / 2 = 6.83333

N

Returns the number of nonmissing observations in a column.

NMISS

Returns the number of missing observations in a column.

COUNT

Returns the total number of observations in a column.

When used as an option with TALLY on page 973 (**Stat** > **Tables** > **Tally** > **Tally Individual Variables**) or the STATS on page 144, COUNT computes the number of observations in each group. For example, if you tally a column containing 1s, 2s and 3s, COUNT tells you how many of each there are. Using COUNTS with STATS, you can store these counts, which is useful if you want to generate a frequency plot.

CUMN

CUMN, or cumulative n, computes a cumulative frequency count of the number of non-missing observations in each group listed in the BY columns. If there are no BY columns listed, CUMN simply counts the non-missing observations in the columns listed with DESCRIBE or STATS. CUMN counts observations for non-missing groups only, unless you include the MISSING subcommand.

PERCENT

If you want statistics for different groups listed in one or more BY columns, PERCENT computes what percentage of the whole is accounted for by each group.

For example, suppose you list one BY column containing five 1s, one 2, and four 3s. PERCENT computes 50% for the first group, 10% for the second, and 40% for the third.

If you omit the BY subcommand, then PERCENT and CUMPERCENT calculate the value 100 (for 100%).

If you include columns on the main DESCRIBE or STAT command line, then PERCENT and CUMPERCENT calculate percentages for only the non-missing values in those columns.

CUMPERCENT

If you want statistics for different groups listed in one or more BY columns, CUMPERCENT computes the cumulative percentage represented by each group.

For example, suppose you list one BY column containing five 1s, one 2, and four 3s.

If you omit the BY subcommand, then PERCENT and CUMPERCENT calculate the value 100 (for 100%).

If you include columns on the main DESCRIBE or STAT command line, then PERCENT and CUMPERCENT calculate percentages for only the non-missing values in those columns.

Graphs

Each graph subcommand displays a separate graph for each variable listed on the command line.

If you use the BY subcommand, Minitab generates one graph for each column listed on the command line.

GHISTOGRAM

Displays a histogram for each variable.

GHISTOGRAM and GNHISTOGRAM display a histogram for each level of the BY variable in a separate panel of the graph. The histograms for one variable are on the same scale to facilitate comparisons between the levels.

GNHISTOGRAM

Displays a histogram with a normal curve for each variable. GNHISTOGRAM bases each normal curve on the sample mean and standard deviation of the data in the corresponding histogram.

GHISTOGRAM and GNHISTOGRAM display a histogram for each level of the BY variable in a separate panel of the graph. The histograms for one variable are on the same scale to facilitate comparisons between the levels.

GINDPLOT

Displays an individual value plot for each variable.

GINDPLOT and GBOXPLOT display side-by-side plots in a single graph, one for each level of the BY variable. This allows you to easily compare the different levels of the BY variable.

GBOXPLOT

Displays a boxplot for each variable.

GINDPLOT and GBOXPLOT display side-by-side plots in a single graph, one for each level of the BY variable. This allows you to easily compare the different levels of the BY variable.

GSUMMARY

Displays a graphical summary for each variable.

Quartiles

Every group of data has three quartiles. To calculate quartiles, use Sort (**Data** > **Sort**) to order the data from smallest to largest. The first quartile (Q1) is the observation at position (n+1) / 4. The second quartile is the median. The third quartile (Q3) is the observation at position 3(n+1) / 4, where n is the number of observations. If the position is not an integer, interpolation is used.

For example, suppose n=10. Then (10 + 1)/4 = 2.75, and Q1 is between the second and third observations (call them x2 and x3), three-fourths of the way up. Thus, Q1 = x2 + 0.75(x3 – x2). Since 3(10 + 1)/4 = 8.25, Q3 = x8 + 0.25(x9 – x8), where x8 and x9 are the eight and ninth observations.

QONE

The first quartile is also referred to as the 25^{th} percentile because data from 25% of the observations are less than or equal to this value.

IQRANGE

The interquartile range equals Q3 – Q1.

QTHREE

The third quartile is also referred to as the 75^{th} percentile because data from 75% of the observations are less than or equal to this value.

STATS: Session command for storing descriptive statistics

STATS [C...C]

STATS computes a wide range of statistics on entire columns or subsets of columns and stores those statistics in the worksheet.

STATS is similar to the commands TALLY on page 973, TABLE on page 971, and DESCRIBE on page 140. TALLY, TABLE, and DESCRIBE compute many of these statistics and display them in an easy-to-read table format. But TABLE and DESCRIBE cannot store the values of these statistics for use in further analysis as STATS can.

Here is an example using the PULSE.MTW data set. This data set includes three columns named Height, Weight, and Sex (containing 1's for males, 2's for females). The NAME command names the three new columns that STATS will store its results in. GVALUES will store the numbers 1 and 2 into C11, because these are the distinct values in the BY column named Sex. MEAN will store the mean height for males and the mean height for females in C12, and the mean weight for males and the mean weight for females in C13.

```
NAME C11 'SexID' C12 'MeanHt' C13 'MeanWt'
STATS 'Height' 'Weight';
 BY 'Sex';
 GVALUES C11;
 MEAN C12 C13.
```

When there is no BY subcommand, you must list at least one column on the main STATS command line. In that case, STATS computes the requested statistics for the entire column.

When you omit columns from the main STATS command line, you must use the BY subcommand. In that case, you may store these statistics: N, NMISS, COUNT, CUMN, PERCENT, and CUMPERCENT. For more information, go to Notes on subcommands that store descriptive statistics (STATS command) on page 1255.

You can use either the GLABELS, GVALUES, or GIDS subcommand so that you know which group each row of statistics belongs to. Use the NAME on page 56 command to name each column so that you know what you stored.

Control of grouping

BY C...C

BY lists the columns that contain the group variables (such as a column named Temp containing the values Low, Medium, and High). Columns listed with BY may contain numeric, text, or date/time data. When you include the BY subcommand, STATS computes statistics for each group listed in the BY columns or columns. When you omit the BY subcommand, STATS computes statistics for whole columns rather than for subgroups.

NOEMPTY

NOEMPTY omits empty cells.

MISSINGS

MISSINGS includes missing as a distinct value of the BY column (when the BY column includes a missing value). Otherwise, rows with missing values in BY columns are omitted.

Options

These subcommands store each distinct value listed in each BY column. Therefore, use the same number of columns as you list with BY. It is a good idea to include at least one of these subcommands so you will know which group each row of statistics belongs to.

GLABELS C...C

GLABELS stores the actual labels from your BY columns (such as 10, 20, 30, or Low, Medium, High). GLABELS always stores the labels as text.

GVALUES C...C

GVALUES also stores the actual labels from your BY columns, but it stores them as the same data type (numeric or text) as the BY variable.

GIDS C...C

GIDS stores a 1 for the first value in your BY column, a 2 for the second value, and so on, in a numeric column. If the BY column is numeric, GIDS stores a 1 for the smallest number in the BY column, a 2 for the next largest number, and so on. If the BY column contains text, GIDS stores a 1 for the first entry (row 1), a 2 for the next entry that is different, and so on.

EXPAND

Stores a row of the desired statistic for each row of input instead. If you do not issue EXPAND, Minitab stores the statistic in the first row.

Statistics

MEAN

Stores the arithmetic mean, or average. The mean is a commonly used measure of the center of a batch of numbers.

Note Missing values are omitted from the calculation of the function Mean.

SEMEAN

Stores the standard error of the mean. It is calculated as StDev / SQRT (N).

STDEV

Stores the sample standard deviation, which provides a measure of how spread out the data are. To calculate the variance, simply square the standard deviation value.

If the column contains x1, x2, ..., xn, with mean $\bar{x}$, then the standard deviation is calculated as follows:

$$\sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

Note Missing values are omitted from the calculation of the function Standard deviation.

VARIANCE

Stores the sample variance, which is a measure of how far the data are spread about the mean. Sample variance equals the standard deviation squared. It can also be computed with this formula:

$$\frac{\Sigma(x-\bar{x})^2}{n-1}$$

CVARIATION

Stores the coefficient of variation, a measure of relative variability. It is calculated as: 100 (s / $\bar{x}$).

MEDIAN

Stores the median of a column. The median is in the middle of the data: half the observations are less than or equal to it. Suppose the column contains n values. If n is odd, the median is the value in the middle. If n is even, the median is the average of the two middle values.

Note Missing values are omitted from the calculation of the function Median.

TRMEAN

Stores a 5% trimmed mean. Minitab removes the smallest 5% and the largest 5% of the values (rounded to the nearest integer).

SUMS

Stores the sum.

Note Missing values are omitted from the calculation of the function Sum.

MIN

Stores the smallest number in a column.

Note Missing values are omitted from the calculation of the function MIN.

MAX

Stores the largest number in a column.

Note Missing values are omitted from the calculation of the function MAX.

RANGE

Stores the difference between the largest and smallest data value.

SSQ

Stores the sum of squares. Squares each value in the column, and computes the sum of those squared values. That is, if the column contains x_1, x_2, ..., x_n, then sum of squares calculates ($x_1^2 + x_2^2 + ... + x_n^2$).

Note Missing values are omitted from the calculation of the function the function Sum of Squares (corrected).

SKEWNESS

Stores the skewness value. Skewness is a measure of asymmetry. A value more than or less than zero indicates skewness in the data. But a zero value does not necessarily indicate symmetry.

KURTOSIS

Stores the kurtosis value. Kurtosis is one measure of how different a distribution is from the normal distribution. A positive value characterizes a distribution with a heavier tails than the normal distribution. A negative value characterizes a distribution with lighter tails than the normal distribution.

MSSD

Stores half the Mean of the Squared Successive Differences of a batch of numbers. For example, suppose a column contains 1, 2, 4, and 10. The successive differences are 2 – 1 = 1, 4 – 2 = 2, and 10 – 4 = 6. Then:

MSSD = mean(1-squared, 2-squared, and 6-squared) / 2 = 6.83333

N

Stores the number of nonmissing observations in a column.

NMISS

Stores the number of missing observations in a column.

COUNT

Stores the total number of observations in a column.

When used as an option with TALLY on page 973 (**Stat** > **Tables** > **Tally** > **Tally Individual Variables**) or the STATS command, COUNT computes the number of observations in each group. For example, if you tally a column containing 1s, 2s and 3s, COUNT tells you how many of each there are. Using COUNTS with STATS, you can store these counts, which is useful if you want to generate a frequency plot.

CUMN

Stores a cumulative frequency count of the number of non-missing observations in each group listed in the BY columns. If there are no BY columns listed, CUMN simply counts the non-missing observations in the columns listed with DESCRIBE or STATS. CUMN counts observations for non-missing groups only, unless you include the MISSING subcommand.

PERCENT

If you want statistics for different groups listed in one or more BY columns, PERCENT computes what percentage of the whole is accounted for by each group.

For example, suppose you list one BY column containing five 1s, one 2, and four 3s. PERCENT computes 50% for the first group, 10% for the second, and 40% for the third.

If you omit the BY subcommand, then PERCENT and CUMPERCENT calculate the value 100 (for 100%).

If you include columns on the main DESCRIBE or STAT command line, then PERCENT and CUMPERCENT calculate percentages for only the non-missing values in those columns.

CUMPERCENT

If you want statistics for different groups listed in one or more BY columns, CUMPERCENT computes the cumulative percentage represented by each group.

For example, suppose you list one BY column containing five 1s, one 2, and four 3s.

If you omit the BY subcommand, then PERCENT and CUMPERCENT calculate the value 100 (for 100%).

If you include columns on the main DESCRIBE or STAT command line, then PERCENT and CUMPERCENT calculate percentages for only the non-missing values in those columns.

GSUMMARY: Session command for displaying a graphical summary of each variable

GSUMMARY

Displays a graphical summary of each variable.

BY

Lists the columns that contain the group variables (such as a column named Temp containing the values Low, Medium, and High). Columns listed with BY may contain numeric, text, or date/time data.

When you include the BY subcommand, GSUMMARY creates a summary for each group listed in the BY column or columns. When you omit the BY subcommand, GSUMMARY creates a summary for whole columns rather than for subgroups.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

SMCONF C C
Stores the confidence interval for the median in C and C.

SSCONF C C
Stores the confidence interval for the standard deviation in C and C.

ONEZ: Session command for performing a 1-sample Z-test

ONEZ C...C

ONEZ K K
Performs a 1-sample Z-test for the raw or summarized data. You need to provide a value for population standard deviation the SIGMA subcommand. When you do not know the population standard deviation, you can use ONET on page 149.

ONEZ C...C performs a Z-test for each of the columns in C...C.

ONEZ K K performs a Z-test using summarized data with sample size and sample mean in K and K, respectively.

SIGMA K (required)
Specifies the value of the population standard deviation.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

TEST K
Tests the null hypothesis using the hypothesized mean specified in K.

SPVALUE C
SPVALUE stores the p-value of the test in a column.

SCONF C C
SCONF stores the confidence interval for the mean in two columns.

Graphs

Displays a graph for each column listed with ONEZ. Each graph displays the sample mean and a K% confidence interval for the mean. No graph is displayed for the summarized data.

GHISTOGRAM
Displays a histogram.

GINDPLOT
Displays an individual value plot.

GBOXPLOT
Displays a boxplot.

ONET: Session command for performing a 1-sample t-test

ONET C...C

ONET K K K
Performs a t-test using raw or summarized data. The population standard deviation for the test is estimated form the columns listed with ONET. A test value is specified with the TEST subcommand. ONET performs a two-sided test unless you use the ALTERNATIVE subcommand to specify a one-sided test. If you know the population standard deviation or deviations, use ONEZ on page 148 instead.

ONET C...C performs a t-test for each of the columns in C...C.

ONET K K K performs a t-test for the summarized data using sample size, sample mean, and population standard deviation specified in K, K, and K, respectively.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

TEST K
Tests the null hypothesis using the hypothesized mean specified in K.

SPVALUE C
SPVALUE stores the p-value of the test in a column.

SCONF C C
SCONF stores the confidence interval for the mean in two columns.

Graphs

Displays a graph for each column listed with ONET. Each graph displays the sample mean and a K% confidence interval for the mean. No graph is displayed for the summarized data.

GHISTOGRAM
Displays a histogram.

GINDPLOT
Displays an individual value plot.

GBOXPLOT
Displays a boxplot.

TWOT: Session command for performing a 2-sample t-test when samples are in one column

TWOT C C

TWOT K K K K K K
TWOT performs a two-sample t-test and confidence interval when the samples and subscripts are in separate columns or when you have summarized data. Use TWOSAMPLE on page 151 when the samples are in different columns.

You can specify a confidence level to change the default level for the confidence interval produced by TWOT. The default level is 95%. K can be any number between 1 and 100. For example, if you enter the command TWOT 90 C1 C2, Minitab calculates a 90% confidence interval.

TWOT C C performs a 2-sample t-test when the samples are in one column and subscripts are in another column.

TWOT K K K K K K performs a 2-sample t-test using summarized data with sample size, mean, and standard deviation listed for each sample.

ALTERNATIVE K
Specifies a one-sided test. K = –1 gives H1: μ < K, and K = +1 gives H1: μ > K1.

TEST K
Specifies the null hypothesis value K.

POOLED
Uses a pooled procedure to estimate σ. This procedure assumes the two populations have equal variances.

The POOLED procedure is slightly more powerful than the method that does not assume equal variances, but can be seriously in error if the variances are not equal. Thus, the POOLED subcommand should not be used in most cases.

GINDPLOT
The graph for GINDPLOT contains two individual value plots, one for each sample. The individual value plots displays the sample mean and a K% confidence interval for each sample.

GBOXPLOT
The graph for GBOXPLOT contains two boxplots, one for each sample. The boxplots display the sample mean and a K% confidence interval for each sample.

SPVALUE C
SPVALUE stores the p-value of the test in a column.

SCONF C C
SCONF stores the confidence interval for the mean in two columns.

TWOSAMPLE: Session command for performing a 2-sample t-test when the samples are in different columns

TWOSAMPLE
TWOSAMPLE performs a 2-sample t-test and confidence interval when the samples are in different columns. Use TWOT on page 150 when samples and subscripts are in separate columns or when you have summarized data.

You can specify a confidence level to change the default level for the confidence interval produced by TWOSAMPLE. The default level is 95%. K can be any number between 1 and 100. For example, if you enter the command TWOSAMPLE 90 C1 C2, then a 90% confidence interval is calculated.

ALTERNATIVE K
Specifies a one-sided test. K = –1 gives H1: $\mu < K$, and K = +1 gives H1: $\mu > K1$.

TEST K
Specifies the null hypothesis value K.

POOLED
Uses a pooled procedure to estimate σ. This procedure assumes the two populations have equal variances.

The POOLED procedure is slightly more powerful than the method that does not assume equal variances, but can be seriously in error if the variances are not equal. Thus, the POOLED subcommand should not be used in most cases.

GINDPLOT
The graph for GINDPLOT contains two individual value plots, one for each sample. The individual value plots displays the sample mean and a K% confidence interval for each sample.

GBOXPLOT
The graph for GBOXPLOT contains two boxplots, one for each sample. The boxplots display the sample mean and a K% confidence interval for each sample.

SPVALUE C
SPVALUE stores the p-value of the test in a column.

SCONF C C
SCONF stores the confidence interval for the mean in two columns.

PAIR: Session command for performing a paired t-test

PAIR C C

PAIR K K K
Performs a paired t-test. This test is appropriate for testing the mean difference between paired observations when the paired differences follow a normal distribution.

Use to calculate a confidence interval and perform a hypothesis test of the mean difference between paired observations in the population. A paired t-procedure matches responses that are dependent or related in a pairwise manner. This matching allows you to account for variability between the pairs usually resulting in a smaller error term, thus increasing the sensitivity of the hypothesis test or confidence interval.

Typical examples of paired data include measurements on twins or before-and-after measurements. For a paired t-test, the hypotheses are as follows:

H_0: $\mu_d = \mu_0$ versus H_1: $\mu_d \neq \mu_0$

where μ_d is the population mean of the differences and μ_0 is the hypothesized mean of the differences.

When the samples are drawn independently from two populations, use TWOSAMPLE on page 151 or TWOT on page 150.

PAIR C C performs a paired t-test when the data are in columns.

PAIR K K K performs a paired t-test using summarized data when K = sample size, K = mean difference, and K = standard deviation difference.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K
Specifies a one-sided test. K = –1 gives H_1: $\mu_d < \mu_0$, and K = +1 gives H_1: $\mu_d > \mu_0$.

TEST K
Conducts a test of the null hypothesis $\mu = K$.

SPVALUE C
SPVALUE stores the p-value of the test in a column.

SCONF C C
SCONF stores the confidence interval for the mean in two columns.

Graphs

Each of the following graphs displays the sample mean, a K% confidence interval for the mean, and the value under H_0.

GHISTOGRAM
Displays a histogram of the paired differences.

GINDPLOT
Displays an individual value plot of the paired differences.

GBOXPLOT
Displays a boxplot of the paired differences.

PONE: Session command for performing a hypothesis test of the proportion

PONE C...C

PONE K K...K
Performs a test of one binomial proportion.

Use PONE to compute a confidence interval and perform a hypothesis test of the proportion. For example, an automotive parts manufacturer claims that his spark plugs are less than 2% defective. You could take a random sample of spark plugs and determine whether or not the actual proportion defective is consistent with the claim. For a two-tailed test of a proportion:

H_0: $p = p_0$ versus H_1: $p \neq p_0$

where p is the population proportion and p_0 is the hypothesized value.

To compare two proportions, use PTWO on page 155.

PONE C...C performs an analysis of one proportion on samples in columns.

PONE K K...K performs an analysis of one proportion on summarized data where K = the number of trials and subsequent Ks = the number of events.

EVENT K
Specifies the event of interest in a column of data. Use a value from the column for K. Enclose text values in quotation marks. For example, to assess the proportion of "Failure" in a column, enter EVENT "Failure".

TEST K
Specifies the null hypothesis value K.

Options

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K

Specifies the alternative hypothesis and the type of confidence interval. K = –1 gives H1: $p < p_0$ and an upper bound for p. K = +1 gives H1: $p > p_0$ and a lower bound for p. The default of K = 0 gives H1: $p \neq p_0$ and a two-sided interval for p.

ABLAKER

When the argument for ALTERNATIVE is 0, specifies to use the adjusted Blaker's exact method for calculating the two-sided hypothesis test and a two-sided confidence interval.

When the argument for the ALTERNATIVE subcommand is -1 or 1, specifies to use the Clopper-Pearson method to calculate one-sided hypothesis tests and confidence intervals.

CTOLERANCE K

Specifies the convergence criterion for the calculation of the p-value and confidence interval for the adjusted Blaker's exact method. K is a positive, real number. The default value is 10^{-10}.

ITERATION K

Specifies the maximum number of iterations for the calculation of the p-value and confidence interval for the adjusted Blaker's exact method. K is a non-negative integer. The default value is 100.

WILSON

Specifies to use the classical Wilson-score approximate procedure for calculating the hypothesis test and confidence interval. The method is also known as the score method. To use the method with a continuity correction, issue the CCORRRECTION sub-subcommand.

CCORRRECTION

Specifies to use the Wilson-score approximate method with a continuity correction. Without a continuity correction, the method is liberal for small to moderate sample sizes. The continuity correction makes the actual confidence level and the actual alpha value at least the levels that the analysis specifies.

ACOULL

Specifies to use the Agresti-Coull approximate procedure for calculating the hypothesis test and confidence interval.

CPEARSON

Specifies to use the Clopper-Pearson exact method for calculating the hypothesis test and confidence interval.

USEZ

Specifies to use the normal approximation to the binomial distribution for calculating the hypothesis test and confidence interval.

Storage

SPVALUE C

SPVALUE stores the p-value of the test in a column.

SCONF C (C)

For the default analysis or when the argument for ALTERNATIVE is 0, stores the lower side of the confidence interval in the first column and the upper side of the confidence interval in the second column.

When the argument for ALTERNATIVE is -1 or 1, stores the confidence limit in one column.

PTWO: Session command for performing a hypothesis test of the difference between two proportions

PTWO C C

PTWO K K K K

Performs a test of two binomial proportions.

Use PTWO to compute a confidence interval and perform a hypothesis test of the difference between two proportions. For example, suppose you wanted to know whether the proportion of consumers who return a survey could be increased by providing an incentive such as a product sample. You might include the product sample with half of your mailings and see if you have more responses from the group that received the sample than from those who did not. For a two-tailed test of two proportions:

H_0: $p_1 = p_2$ versus H_1: $p_1 \neq p_2$

where p_1 and p_2 are the proportions of success in populations 1 and 2, respectively.

To test one proportion, use PONE on page 153.

You can input either raw or summarized data.

Raw data can be entered in two ways: stacked and unstacked.

- Enter both samples in a single column (stacked) with a group column to identify the population. Columns may be numeric, text, or date/time. Successes and failures are determined by numeric or alphabetical order. Minitab defines the lowest value as the failure; the highest value as the success. For example:
  - For the numeric column entries of "5" and "10", observations of 5 are considered failures; observations of 10 are considered successes.
  - For the text column entries of "agree" and "disagree," observations of agree are considered failures; observations of disagree are considered successes. If the data entries are "yes" and "no", observations of no are considered failures; observations of yes are considered successes.
- Enter each sample (unstacked) in separate numeric or text columns. Both columns must be the same type-numeric or text. Successes and failures are defined as above for stacked data.

You can reverse the definition of success and failure in a text column by applying a value order. For more information, go to VORDER on page 138.

The sample sizes do not need to be equal. Minitab automatically omits missing data from the calculations.

For raw data, enter the number of trials and the number of successes for each sample on the main command line. Enter four integers: the number of trials and the number of successes in the first sample followed by the number of trials and the number of successes in the second sample.

PTWO C C performs a test of two proportions on samples in columns.

PTWO K K K K performs a test of two proportions on summarized data where K = the number of trials and K = the number of events for the first proportion, and K = the number of trials and K = the number of events for the second proportion.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K
Specify a one-sided test. K = –1 gives H1: p < K, and K = +1 gives H1: p > K1.

TEST K
Specifies the null hypothesis value K.

STACKED
Specifies that the data have been entered in a stacked format: the raw data is in one column and the subscripts or group column to identify the population are in a second column.

POOLED
Specifies to use a pooled estimate of p to calculate the test statistic.

SPVALUE C
Stores the p-value of the test in a column.

SFISHER C
Stores the p-value for the Fisher's exact test in a column.

SCONFC C
Stores the confidence interval for the mean in two columns.

ONERATE: Session command for performing a 1-sample Poisson rate test

ONERATE C...C

ONERATE K K...K
Performs a hypothesis test and calculates a confidence interval for the population Poisson rate for each sample you input. You can enter data in raw, summarized, or frequency format. Specify a hypothesized test value with the TEST subcommand. Minitab performs a two-sided test unless you use the ALTE subcommand to specify a one-sided test. Specify the length of the observation space with the LENGTH subcommand to analyze the mean number of occurrences in addition to the occurrence rate.

ONERATE C...C calculates a confidence interval for the population Poisson rate for each column of raw sample data.

ONERATE K K...K calculates a confidence interval for the population Poisson rate for summarized data. The first K equals the sample size; subsequent values of K correspond to the number of events in each summarized sample. You cannot use this command to analyze samples with different sample sizes.

FREQ C...C
If your data exists in frequency format, use this subcommand to specify the frequency columns. Enter them in the same order in which you enter the columns of unique observations.

LENGTH K [K]...[K]
Specify the length of the observation space for each sample to obtain confidence intervals of mean number occurrences in addition to the rate of occurrences. By default, length equals 1, and the occurrence rate equals the mean number of occurrences.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

TEST K

Performs a hypothesis test of population Poisson rate. Enter the hypothesized value of the population Poisson rate for K.

USEZ

Specifies the normal approximation method for calculating the hypothesis test and confidence interval. By default, Minitab uses the exact method instead of the normal approximation method. This subcommand accepts no arguments.

TWORATE: Session command for performing a 2-sample Poisson rate test

TWORATE C C

TWORATE K K K K

Use TWORATE C C to perform a hypothesis test and calculates a confidence interval for the difference between the Poisson rates of two populations. For unstacked data, the arguments correspond to the two columns of data. For stacked data, the first C corresponds to the column of stacked sample data, and the second C corresponds to the column of subscripts.

Use TWORATE K K K K with summarized data to perform a hypothesis test and calculate a confidence interval for the difference between the Poisson rates of two populations. The first K equals the size of the first sample; the second K equals the number of occurrences in the first sample; the third K equals the size of the second sample; the fourth K equals the number of occurrences in the second sample.

LENGTH K [K]

Specify the length of observation for each sample to analyze both the mean number of occurrences and the occurrence rate. If you enter one K, this value becomes the length for both samples. If you enter two values for K, the first becomes the length for the first sample, and the second K becomes the length of the second sample.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

TEST K

Specifies the hypothesized value for the difference between the two population Poisson rates. The default value is 0.

STACKED

Indicates that the data for both samples are stacked in a single column. This subcommand accepts no arguments.

FREQ C [C]

Use this subcommand if your data exists in frequency format. For stacked data, enter only one column of frequencies. For unstacked data, enter two columns of frequency data in the same order in which you enter the columns of unique observations.

POOLED

Uses a pooled estimate of the occurrence rate for both populations. The argument for the TEST subcommand must equal zero to use this option. By default, Minitab estimates the rate of each population separately by using the observed sample rates. This subcommand accepts no arguments.

ONEV: Session command for performing a 1 variance test

ONEV C...C

ONEV K K

For each sample you input, this command performs a one variance hypothesis test and produces a confidence interval for the population variance. A test value is specified with the STEST or VTEST subcommands. Minitab performs a two-sided test unless you use the ALTERNATIVE subcommand to specify a one-sided test. If you prefer to work in terms of standard deviation instead of variance, use the STDEV subcommand.

ONEV C...C calculates a confidence interval for the population variance for each column of raw data. Output includes results of the chi-square method (for normal distributions) and the Bonett method (for any continuous distribution).

ONEV K K calculates a confidence interval for the population variance with summarized data. The first K equals the sample size; the second K equals the sample variance. Output only includes results of the chi-square method (for normal distributions).

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

STEST K

STEST performs a hypothesis test of the population standard deviation.

STEST and VTEST are mutually exclusive.

VTEST K

VTEST performs a hypothesis test of the population variance. K specifies the hypothesized value of the population and must be a positive number.

STEST and VTEST are mutually exclusive.

STDEV

This subcommand indicates that input values refer to standard deviation instead of variance, where applicable. Furthermore, this subcommand causes Minitab to produce output in terms of standard deviation instead of variance. By default, Minitab interprets input and produces output in terms of variance. This subcommand accepts no arguments.

TWOVARIANCES: Session command for determining whether the variances or standard deviations of two groups differ

TWOVARIANCES performs hypothesis tests and computes confidence intervals for the ratios between two populations' variances and standard deviations. You can use this command when samples and subscripts are in separate columns, when each sample is in a single column, or when you have summarized data.

TWOVARIANCES C C

Performs two variances test when sample data are in different columns.

TWOVARIANCES C C

Performs two variances test when all sample data are in one column C and subscripts are in a second column C. Use STACKED with this option.

STACKED

Specifies that data are stacked.

TWOVARIANCES K K K K

Performs two variances test using summarized data. Specify the sample size and variance for Sample 1 and the sample size and variance for Sample 2. If you specify standard deviations instead of variances, use the STDEV subcommand.

Options

STDEV

Indicates that the input summary values refer to standard deviations instead of variances.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

STEST K

STEST specifies the hypothesized value K for the ratio of two standard deviations. K must be a positive number. The default is 1, which tests for equal standard deviations. STEST and VTEST are mutually exclusive.

VTEST K

VTEST specifies the hypothesized value K for the ratio of two variances. K must be a positive number. The default is 1, which tests for equal variances. STEST and VTEST are mutually exclusive.

SNAMES K K

Specifies names for the samples when you enter summarized data. The default sample names are "Sample 1" and "Sample 2".

USEF

Specifies to use the F-test method instead of Bonett's method and Levene's method. The F-test is accurate only for normally distributed data. Any departure from normality can cause the F-test to yield inaccurate results. However, if the data conform to the normal distribution, then the F-test is typically more powerful than either Bonett's test or Levene's test.

Graphs

GINTERVAL

GINTERVAL displays a graphical summary of the results including the following:

- Confidence intervals for the ratio of the standard deviations or variances
- P-values for the hypothesis tests
- Boxplots of each sample

GHISTOGRAM
GHISTOGRAM displays a graph containing a histogram for each sample.

GINDPLOT
GINDPLOT displays a graph containing an individual value plot for each sample.

Results

NODEFAULT
Specifies that none of the following results are displayed.

TMETHOD
TMETHOD displays the method table which includes the null hypothesis, the alternative hypothesis, and the significance level (denoted by α or alpha).

TSTATISTICS
TSTATISTICS displays the statistic table which includes the standard deviation of each sample, the variance of each sample, and the confidence interval for the standard deviation or the variance of each sample. The ratio of the standard deviations and the ratio of the variances are also displayed.

TCONFIDENCE
TCONFIDENCE displays the confidence interval table which includes confidence intervals for the ratio of the standard deviations and the ratio of the variances.

TTEST
TTEST displays the test table which includes the degrees of freedom, the test statistics, and the p-values. When the sample sizes are unequal, the test statistic for Bonett's method is undefined. However, the p-value can be calculated by inverting the confidence interval procedure.

CORRELATION: Session command for measuring the strength and direction of the association between two variables

CORRELATION C...C
Calculates the Pearson product moment correlation or the Spearman rank-order correlation for every pair of samples in C...C.

Options

Minitab calculates the Pearson product moment correlation or the Spearman rank-order correlation and the associated confidence intervals and p-values.

PEARSON
Calculates the Pearson (product-moment) correlation. Use the Pearson correlation coefficient to examine the strength and direction of the linear relationship between two continuous variables.

For each pair of samples, Minitab calculates the p-value for the hypothesis test of zero-correlation and the confidence interval for the Pearson correlation. The confidence intervals are based on the Fisher's Z transformation.

SPEARMAN

Calculates the Spearman rank-order correlation (also called Spearman's rho). The Spearman correlation evaluates the monotonic relationship between two continuous or ordinal variables. In a monotonic relationship, the variables tend to change together, but not necessarily at a constant rate. The Spearman correlation coefficient is based on the ranked values for each variable rather than the raw data.

Spearman correlation is often used to evaluate relationships involving ordinal variables. For example, you might use a Spearman correlation to evaluate whether the order in which employees complete a test exercise is related to the number of months they have been employed.

For each pair of samples, Minitab calculates the p-value for the hypothesis test of zero-correlation and the confidence interval for the Spearman correlation. The confidence intervals are based on the Fisher's Z transformation of the Spearman correlation and Bonett and Wright adjustments.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter CONFIDENCE 90. The default value of K is 95.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table which includes the correlation method. When the matrix plot includes p-values or the output includes the table of pairwise correlations, the Method table also includes the hypotheses and the significance level.

TCORRELATION

Displays the correlation matrix.

TPCORRELATION *M*

Displays the table of pairwise correlations. The table includes pairwise sample sizes, the correlation estimates, the confidence intervals, and the p-values.

Graphs

GMPLOT

Displays a matrix of scatterplots for all pairs. By default, the matrix plot includes correlation values and the matrix plot is the lower left triangle of the matrix so that each pair of variables is on one plot. There are two groups of optional subcommands used to control the statistics and the display of the plot.

Use one of the following optional subcommands to change the display statistics:

R

Displays the correlation values on the plots.

RPVALUES

Displays the correlation values and the p-values for the test of zero-correlation on the plots.

RCIS
Displays the correlation values and the confidence intervals on the plots.

NOSTATS
Suppresses the display of statistics on the plots.

Use one of the following optional subcommands to specify how to display the matrix plot:

LL
Displays the lower left triangle portion of the matrix plot.

UR
Displays the upper right triangle portion of the matrix plot.

FULL
Displays the entire matrix plot (both the lower left triangle and the upper right triangle) of the matrix plot.

Storage

SCORRELATION *M*
Stores the correlation matrix.

SPVALUES *M*
Stores a matrix that contains the p-values for the hypothesis test of zero-correlation.

SCIS *M M*
Stores two matrices that contain the lower bounds and the upper bounds of the confidence intervals for the pairwise correlations. The first matrix stores the lower bounds and the second matrix stores the upper bound.

COVARIANCE: Session command for calculating the covariance between pairs of columns

COVARIANCE C...C *[M]*
Calculates the covariance between each pair of columns. You can display the result or you can store the covariance matrix.

When you list two columns on the command line, Minitab calculates the covariance for the pair. When you list more than two columns, Minitab calculates the covariance for every pair of columns, and displays the lower triangle of the resulting covariance matrix (in blocks if there is insufficient room to fit across a page). Minitab does not display the covariance matrix when the matrix is stored in M.

NORMTEST: Session command for performing a normality test

NORMTEST C

Generates a normal probability plot.

Normal plots use the values in the input column as x-values. The grid on the graph resembles the grids found on normal probability paper. The horizontal axis is a linear scale. The line forms an estimate of the cumulative distribution function for the population from which data are drawn.

By default, an Anderson-Darling test for normality is performed and the numerical results are displayed with the graph. You can also use a Ryan-Joiner test (similar to a Shapiro-Wilk test) or a Kolmogorov-Smirnov test.

PTILES C

PTILES K...K

Specifies a set of reference percents. The values must be between 0 and 100 when percents are used as the y-scale type or 0 to 1 when probability is the y-scale type. Minitab marks each percent in the column with a horizontal reference line on the plot, and marks each line with the percent value. Minitab draws a vertical reference line where the horizontal reference line intersects the line fit to the data, and marks this line with the estimated data value.

DVALUE

Use DVALUE to show the percents at the reference x-scale positions specified in PTILES.

PERCENT

Specifies a percent y-scale.

PROBABILITY

Specifies a probability y-scale.

SCORES

Specifies a percentile y-scale.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

SPVALUE C

SPVALUE stores the p-value of the test in a column.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different

file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Goodness-of-fit tests

There are 3 types of goodness-of-fit test: a chi-square based test, an ECDF based test, and a correlation based test. By default, Minitab uses the Anderson-Darling test, which is an ECDF based test.

When your α-value is greater than the p-value displayed with the graph, you reject the hypothesis of normality. The α-value (also called significance level), is the probability that you will reject the hypothesis of normality when the hypothesis is true.

For example, if you are using an α-value of 0.10 and the p-value is 0.07, then you reject the hypothesis of normality at the 0.10 level.

RJTEST
Use RJTEST to perform a Ryan-Joiner test, which is a correlation based test.

KSTEST
Use KSTEST to perform a Kolmogorov-Smirnov test, which is a chi-square based test.

OUTLIER: Session command for performing an outlier test

OUTLIER C...C

Use to identify a single outlier in a sample. The hypotheses are as follows:

- H_0 (the null hypothesis): All values in the sample are from the same, normally distributed population.
- H_1 (the alternative hypothesis): One of the values in the sample is not from the same, normally distributed population.

You should not use Minitab's outlier tests more than once on the same sample. If you remove an outlier from your sample and then retest, you risk removing values that are not actually outliers.

BY C...C

Lists the columns that contain the grouping variables (such as a column named Temp that contains the values Low, Medium, and High). Columns listed with BY may contain numeric, text, or date time data. When you include the BY subcommand, OUTLIER conducts a separate analysis for each group listed in the BY column or columns.

Options

GRUBBS

GRUBBS specifies to use Grubbs' test, which is the default test.

The arguments K K specify the specific ratio to test. For example, to perform the standard Dixon's Q test, which corresponds to a Dixon ratio of r10, the arguments are 1 and 0. The first argument can either be 1 or 2 and the second argument can be 0, 1, or 2.

DIXON K K

DIXON specifies to use one of the six versions of Dixon's outlier test instead of Grubbs' test.

The arguments K K specify the specific ratio to test. For example, to perform the standard Dixon's Q test, which corresponds to a Dixon ratio of r10, the arguments are 1 and 0. The first argument can either be 1 or 2 and the second argument can be 0, 1, or 2.

ALPHA K

Specifies the significance level (denoted by α or alpha) for the test as K. The default is 0.05.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Graphs

GOUTLIERPLOT

Displays the outlier plot with a summary of the results, including an individual value plot of the data which highlights the outlier if one is identified. The outlier plot also includes summary statistics, the test statistic, and the p-value for the test statistic.

Results

Minitab displays all output tables by default; you do not have to enter the subcommands.

NODEFAULT

If you enter NODEFAULT, then each table is only displayed if you enter TMETHOD, TTEST, or TOUTLIER.

TMETHOD

TMETHOD displays the method table which includes the null hypothesis, the alternative hypothesis, and the significance level significance level (denoted by α or alpha).

TTEST

TTEST displays the test table which includes the specific test performed, summary statistics for the sample, and the p-value for the hypothesis test.

TOUTLIER

TOUTLIER displays the value of the outlier if one is identified, and the row in the worksheet that contains the outlier.

Storage

SINDICATOR C...C

Stores an indicator column for each variable. A one (1) is stored in the row that contains the outlier. A zero (0) is stored in each of the remaining rows.

PGOODNESS: Session command for performing a chi-square goodness-of-fit test for Poisson distribution

PGOODNESS C...C

Performs a chi-square goodness-of-fit test for Poisson distribution for each data set. Each C corresponds to one data set containing nonnegative integers. If the column contains missing data, that row is excluded from all computations.

FREQUENCY C...C

Specifies the frequency column, one for each data set. The length of each column must match the length of the corresponding data set.

MEAN C

MEAN K...K

Specifies the mean of the Poisson distribution for each data set as K...K or in C. If you specify only one value for the mean, then this value applies to all data set.

Graphs

GBAR

Displays a bar chart of the observed and expected values.

GCHISQ

Displays a bar chart of each category's contribution to the chi-square value.

PARETO

Use PARETO with GCHISQ to display a bar chart that orders each category's contribution to the chi-square value from largest to smallest.

Results

RTABLE

Displays a table of the observed values and the expected values and the test results.

Regression

INDICATOR: Session command for creating indicator variables

INDICATOR C C...C

Creates indicator variables (also called dummy variables) that you can use in a regression analysis. If you use REGRESS on page 170, you do not need to create indicator variables.

The smallest number in C2 is 2 and the largest is 6. INDICATOR creates one indicator variable for each unique value.

- C11 is the indicator variable for the value 2. C11 contains a 1 in every row where C2 contains a 2, and 0 otherwise.
- C12 is the indicator variable for the value 3. C12 contains a 1 in every row where C2 contains a 3, and 0 otherwise.
- C13 is the indicator variable for 5. C13 contains a 1 in every row where C2 contains 5, and 0 otherwise.
- C14 is the indicator variable for 6. C13 contains a 1 in every row where C2 contains 6, and 0 otherwise.

If C2 contains an * (missing data code), then all indicator variables are also set to *.

The number of storage columns must be equal to the number of distinct values (not including *) in the input column. Up to 100 storage columns are allowed on INDICATOR.

The following command language is an example of using INDICATOR.

```
INDICATOR  C2, C11-C14
```

Before using INDICATOR

C1
2
2
5
3
6
5
*

After using INDICATOR

C1	...	C11	C12	C13	C14
2	...	1	0	0	0
2	...	1	0	0	0
5	...	0	0	1	0

C1	...	C11	C12	C13	C14
3	...	0	1	0	0
6	...	0	0	0	1
5	...	0	0	1	0
*	...	*	*	*	*

REGRESS: Session command for performing a regression analysis

REGRESS

Performs simple, polynomial, and multiple regression using the least squares method.

REGRESS uses least squares to fit a model to one or more continuous or categorical predictors. Optionally, you can:

- Perform stepwise regression
- Fit the model without an intercept
- Perform weighted regression
- Use a test data set or cross-validation to validate the model
- Store the residuals, fitted values, and many other diagnostics for additional analysis
- Generate point estimates, and prediction and confidence intervals for predicted values
- Transform non-normal data using the Box-Cox transformation
- Generate several plots for residual analysis. For more information, go to Residual analysis and regression diagnostics on page 1261.

Options

Specifies information about the variables and terms to include in the model.

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

CONTINUOUS C...C

Specifies the continuous predictors if you have any. The column(s) must be numeric or date/time and must match the length of the response column.

CATEGORICAL C...C

Specifies the categorical predictors if you have any. The column(s) can be numeric, text, or date/time and must match the length of the response column.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors may be repeated. Nested terms are not allowed. The model can be nonhierarchical.

WEIGHT C

Performs a weighted regression. An n x n matrix W is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X'\ W\ X})^{-1}\ (\mathbf{X'\ W\ Y})$

This is equivalent to minimizing the weighted SS Error,

$\Sigma\ w_i\ (Y - \hat{Y})^2$ where w i is the weight in row i.

CONSTANT

When you use CONSTANT, Minitab includes the β_0 term (the intercept) in the equation. Thus, Minitab fits the model:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e$

NOCONSTANT

When you use NOCONSTANT, Minitab omits the β_0 term (the intercept) from the equation. Thus, Minitab fits the model:

$Y = \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e$

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

TOLER K

Specifies the tolerance level for collinearity and constant check. By default, K = 4 * 2.22e–012.

SSQUARES

Specifies sequential sum of squares for tests in the ANOVA table. The default is the adjusted sums of squares.

EFFECT

Specifies the effect coding (−1, 0, +1) scheme for categorical predictors.

If you do not specify either EFFECT or BINARY, Minitab uses the preference set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

BINARY

Specifies the binary coding (1, 0) scheme for categorical predictors.

If you do not specify either EFFECT or BINARY, Minitab uses the preference set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

REFERENCE C K ... C K

Changes the default coding for the categorical predictor columns. To change the default reference factor level, specify the factor column followed by the reference level. (You must enclose text and date/time levels in double quotation marks.) You can assign a reference level only when you use the binary coding (1, 0) scheme.

Standardizing the continuous predictor

Use this set of subcommands to standardize the continuous predictors in your model. You can use SCALE and LOCATION in conjunction with each other. LEVELS is mutually exclusive with the LOCATION and SCALE subcommands.

If you do not specify LOCATION, SCALE, LEVELS, or UNSTANDARDIZED, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

LOCATION [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by subtracting a constant from each predictor. If you do not specify any arguments, the mean of each predictor column is subtracted. K specifies to subtract a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

SCALE [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by dividing each predictor by a constant. If you do not specify any arguments, each predictor column is divided by the standard deviation. K specifies to divide by a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

LEVELS [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by DOE-type coding for the specified low and high levels K K&K K. The number of arguments must be twice the number of continuous predictors.

UNSTANDARDIZED

Specifies the analysis is to be performed on the original predictors.

Box-Cox

BOXCOX [K]

Performs a Box-Cox transformation with a specified lambda. K is the value of lambda and must be between −5 and +5. If you do not specify K, then Minitab will find the optimal lambda. By default, Minitab rounds the optimal value.

Minitab cannot calculate the optimal lambda for stepwise regression. Consequently, you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

Stepwise

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find an improvement of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

FVALIDATION

Specifies the use of forward selection with the validation R^2 as the criterion for the selection of the final model. When you use `FVALIDATION`, you must also use `TEST` or `CVTEST`. When you use `TEST`, the procedure is similar to forward selection for other methods. At the end of each step, Minitab Statistical Software calculates the test R^2 statistic. At the end of the forward selection procedure, the model with the greatest test R^2 value is the final model.

When you use `CVTEST`, the procedure repeats forward selection on each fold. The procedure evaluates all the folds at each step and identifies the step with the best overall k-fold R^2 value. The last part of the procedure is to perform forward selection on the full dataset, stopping at the best step from the selections on the folds.

`FVALIDATION` stops under the same conditions as `FINFORMATION`. The `ATEND` sub-subcommand that adds terms to the model to make it hierarchical at the end is not compatible with `FVALIDATION`.

GRSQUARE

When you use `GRSQUARE` after `FVALIDATION`, displays a graph of the R^2 statistic and the validation R^2 statistic for each step in the model selection procedure.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Validation

TEST C K

TEST K [K]

Specifies to perform validation with a test data set. When you specify a column you select the rows to include in the test data set. The column has two distinct values. Use the constant to specify which value is for the test data set.

When you specify a constant first, Minitab Statistical Software randomly selects the data for the test data set. The constant is the fraction of the data in the test data set. The optional constant is a seed for the random number generator. If you specify the seed, you can get the same test data set again by specifying the same seed.

CVTEST C

CVTEST K [K]

Specifies to perform k-fold cross validation. When you specify a column, you select the rows for each fold. Rows with the same value in the column are in the same fold.

When you specify a constant, Minitab Statistical Software randomly selects the data for the folds. The constant is the number of folds. The optional constant is a seed for the random number generator. If you specify the seed, you can get the same folds again by specifying the same seed.

Graphs

When you use `TEST`, the residual plots include separate displays for the training data set and the test data set.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of errors. This plot lets you look at both the magnitude and the importance of an effect at the same time. The alpha level is 1– the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram or individual value plot of the residuals, depending on the sample size.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Results

NODEFAULT
Specifies that no default tables or graphs will be displayed

TBASIC
Displays the tables for the Minitab 16 version of REGRESS.

TMETHOD
Displays the method table.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. Use FULL to display the coefficients, p-values, Mallows' Cp, and model summary statistics for each step of the procedure. Use NOFULL to hide these statistics. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Stepwise**.

TEQUATION
Displays the regression equation table. Minitab will display up to 50 equations. If you want to see a single equation, rather than a separate equation for each factor level combination, use SINGLE. To see the separate equations, use SEPARATE. If you do not specify SINGLE or SEPARATE, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

TCOEFFICIENTS
Displays the table of coefficients. Use FULL to display the full set of coefficients for categorical predictors. Use NOFULL to show only the linearly independent coefficients. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

TSUMMARY
Displays the summary of model table.

TANOVA
Displays the ANOVA table.

TDIAGNOSTICS
Displays a table of diagnostics. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations. If you do not specify K, Minitab uses the preference set in **File** > **Options** > **Linear Models** > **Display of Results**.

TDW
Displays Durbin-Watson statistics.

TSIMPLE
Displays the simple versions of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

TEXPAND
Displays the expanded version of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

Storage for Box-Cox transformation

BCRESP C
Stores the Box-Cox transformation of the response in C.

BFITS
Stores the fits for the original response.

Storage for fits and residuals

Use to store the fits and residuals in the specified column. When you use `TEST` , the following statistics are missing for the test data set: studentized residuals, Cook's distance, and DFITS. The Durbin-Watson statistic is for the training data set.

FITS C
Stores the fitted values.

RESIDUALS C
Stores the residuals (fitted values – observed values).

SRESIDUALS C
Stores the standardized residuals.

TRESIDUALS C
Stores the deleted Studentized residuals.

HI C
Stores the leverages.

COOK C
Stores Cook's distance.

DFITS C
Stores the DFITS.

SDW C
Stores the Durbin-Watson statistic.

Storage for characteristics of the estimated equation

Use to store characteristics of the estimated equation in the specified column.

COEFFICIENTS C
Stores the estimated coefficients.

FITS C

Stores the fitted values, often called the Y-hats ($\hat{Y}$).

MSE K

Stores the mean square error. When you use TEST, the statistic is for the training data set.

SPVALUE C

Stores the p-values for each predictor.

SVIF C

Stores the VIFs for each predictor.

SS C

Stores the standard deviation of the error term. When you use TEST, the first row is for the training data set and the second row is for the test data set.

SRSQ C

Stores the R-squared for the model. When you use TEST, the first row is for the training data set and the second row is for the test data set.

SRSADJ C

Stores the adjusted r-squared for the model. When you use TEST, the statistic is for the training data set.

SPRESS C

Stores the PRESS statistic for the model. When you use TEST, the statistic is for the training data set.

XMATRIX *M*

Stores the design matrix for regression model.

XPXINV *M*

Stores a p x p matrix, inverse of **X'X**. This matrix, when multiplied by MSE is the variance-covariance matrix of the coefficients. If the WEIGHTS subcommand is used, then XPXINV stores the inverse of the **X'WX**. When you use TEST, the matrix is for the training data set.

RMATRIX *M*

Stores the R matrix of the QR decomposition, also known as the Cholesky decomposition. When you use TEST, the matrix is for the training data set.

Storage for validation

Use the following commands to store statistics for validation with a test data set or validation with k-fold cross validation.

SSAMPLE C

When CVTEST is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When TEST is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the two values.

CVFIT C
Stores the fitted values from each fold's turn as the test data set.

CVRESID C
Stores the residuals from each fold's turn as the test data set.

BREG: Session command for performing best subsets regression

BREG C C...C
Performs best subsets regression, using the maximum R-squared criterion.

Suppose you specify m predictors. BREG first looks at all one-predictor regression models, selects the model with the largest R-squared, and displays information on this model and the next best one-predictor model. Then BREG looks at all two-predictor models, selects the model with the largest R-squared, and displays information on this and the next best one. The process continues until all m predictors are used.

BREG is an efficient way to select a group of "best subsets" for further analysis by selecting the smallest subset that fulfills certain statistical criteria. The subset model may actually estimate the regression coefficients and predict future responses with smaller variance than the full model using all predictors.

INCLUDE C...C
Includes the predictors specified in C...C in all models. Only columns that were specified with the BREG command can be specified in INCLUDE.

NVARS K [K]
By default, displays the best one-predictor models, the best two-predictor models, on up to the best m-predictor models. If you specify, for example, `NVARS 5 12`, then only the best 5-, 6-, ..., 12-predictor models are displayed.

Note NVARS does not count the predictors given in INCLUDE. Thus, `INCLUDE C1–C3` with `NVARS 2 6`, displays models with 3 + 2 = 5 to 3 + 6 = 9 predictors.

BEST K
Displays information from the "best" K models of each size, if there are that many. K can be from 1 to 5. The default is 2.

NOCONSTANT
When you use NOCONSTANT, Minitab omits the b 0 term (the intercept) from the equation. Thus, Minitab fits the model:

$Y = \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e$

When NOCONSTANT is specified, Minitab does not display R-squared and adjusted R-squared, since their interpretation is difficult.

You can also use NOCONSTANT as a main command. In that case, it applies to all BREG, REGRESS, and STEPWISE commands that follow.

CONSTANT

When you use CONSTANT, Minitab includes the b 0 term (the intercept) in the equation. Thus, Minitab fits the model:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e$

You can also use CONSTANT as a main command. In that case, it applies to all BREG, REGRESS, and STEPWISE commands that follow.

NOWARN

When you specify more than 14 free variables with BREG, Minitab displays a prompt that asks whether you want to continue with BREG computations even though it will take a long time. Use NOWARN if you want to suppress this warning prompt.

TEXPAND

Displays the expanded table that includes PRESS, AICc, BIC, and condition number. You can also set this option in **File** > **Options**.

FITLINE: Session command for creating a fitted line plot

FITLINE C C

Plots the regression line through the actual data or the log 10 of the data. The fitted line plot shows you how closely the actual data lie to the fitted regression line. You can include only one predictor in the model.

Polynomial regression is one method for modeling curvature in the relationship between a response variable (Y) and a predictor variable (X). Polynomial modeling techniques try to account for the curvature by extending the simple linear regression model to include higher powers of X - such as X-squared - as predictors.

Use the POLY subcommand to specify the order of the polynomial you want to fit. If you have not transformed the X or Y variables, you can fit the following models with FITLINE:

Model type	Order	POLY K	Statistical model
Linear	First	1	$Y = \beta_0 + \beta_1 X + e$
Quadratic	Second	2	$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + e$
Cubit	Third	3	$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + e$

You can generate additional models by using the $\log_{10}$ of X and/or Y for linear, quadratic, and cubic models. These models provide another way of modeling curvature. In addition, taking the $\log_{10}$ of Y can reduce right-skewness and some forms of heteroskedasticity, such as unequal variances.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)

Value of K	Type of residuals
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4 ... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Options

POLY K
Specifies the order of the polynomial model you want to fit and plot. K is an integer from 1 to 3. The default is 1.

LOGY
Uses $\log_{10}$ Y as the response variable.

LOGX
Uses $\log_{10}$X as the predictor variable. If LOGX is used with polynomials of order greater than one, then the polynomial regression will be based on powers of the $\log_{10}$X.

YSCALE
Plots the transformed response and predictor variables on a log scale. FITLINE ignores YSCALE if you do not use LOGY.

XSCALE
Plots the transformed response and predictor variables on a log scale. FITLINE ignores XSCALE if you do not use LOGX.

CI
Displays confidence bands about the fitted regression line.

PI
Displays prediction bands about the fitted regression line.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

Storage

RESIDUALS C

Stores the residuals in C. If you specified LOGY, Minitab stores the log-transformed residuals. The residual for the i th log-transformed response is:

$e_i = (\log_{10} Y_i - \log_{10} \hat{Y}_i)$

FITS C

Stores the fitted values, often called the Y-hats ($\hat{Y}$), in C. If you specified LOGY, Minitab stores the log-transformed fits ($\log_{10} \hat{Y}$).

COEFFICIENTS C

Stores the estimated coefficients in C.

BRESIDUALS C

Stores in C the residuals of the transformed response in the original scale of the data. This subcommand is only available if you used LOGY.

BFITS C

Stores in C the fitted values of the transformed response in the original scale of the data. This subcommand is only available if you used LOGY.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB

TIF grayscale

TIF

TIF color

BMPB

BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

SSWORKSHEET: Session command for creating a stability study worksheet

Important SSWORKSHEET writes data to the active worksheet. To prevent possible confusion and lost data, it is best to use NEW immediately prior to SSWORKSHEET. (Do not use NEW however, if you use the TIME subcommand.) NEW creates a new, blank worksheet and makes the new worksheet the active worksheet.

SSWORKSHEET K K K
Creates a stability study worksheet with the following parameters (in the following order):

- The number of times as K
- The number of batches as K
- The number of replicates per batch at each time as K

TIME C

TIME K...K
Optional subcommand TIME specifies the test times as values from column C or as constants K...K.

SSWORKSHEET K K
Creates a stability study worksheet with the following parameters (in the following order):

- The number of batches as K
- The number of replicates per batch at each time as K

TIME C...C (required)
Required subcommand TIME specifies a column of test times for each batch

Options

BNAME C

BNAME K...K
Specifies a name for each batch. The default batch names are the text values "1", "2", "3", ...

TSUMMARY
Displays a summary of the design that includes the number of sampling times, the number of batches, the number of samples per batch at each time, and the total number of runs.

SORDER C
Stores the standard order of the runs in column C.

RORDER C
Stores the run order in column C.

XMATRIX C C
Stores the sampling times for each run in the first column C and the batch names in the second column C.

Randomization

Use the following commands to control how the runs are randomized in the worksheet. If you do not use any of these subcommands, then the runs are not randomized, so the run order is the same as the standard order.

To generate the same random order, you can optionally set the base as K, where K is an integer that is greater than 0.

RANDOMIZE [K] (default)
RANDOMIZE specifies to randomize the order of the batches at each test time, and the repeats at each test time.

BRANDOMIZE [K]
BRANDOMIZE specifies to randomize the order of the batches at each test time, but to keep all repeats from one batch together before switching to the next batch.

NRANDOMIZE
NRANDOMIZE specifies to sample each batch in sequential order and to take all repeats from one batch before switching to the next batch.

SHELFLIFE: Session command for performing a stability study

SHELFLIFE
Analyzes the stability of a product over time to determine the product's shelf life. Minitab fits a linear model to represent the relationship between the response variable, the time variable, and an optional batch factor. The batch factor can be fixed or random. Based on the model, Minitab calculates the shelf life, which is the length of time that the response is expected to remain within the specifications.

Options

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

TIME C (required)

Specifies the time variable. The column must be numeric or date/time and must match the length of the response column.

BATCH C

Specifies the categorical batch factor, if you have one. The column can be numeric, text, or date/time and must match the length of the response column.

RANDOM

Specifies that BATCH is a random factor. If BATCH is not issued, RANDOM is ignored. For random batches, Minitab uses an iterative algorithm to estimate the variance components and to select the model. There can be cases where the variance components cannot be estimated and no further analysis is possible.

MAXITER K

Specifies the maximum number of iterations. K must be a nonnegative integer.

STARTING K K K

Specifies the starting estimates for the variance components. If you do not use STARTING, Minitab uses the minimum norm quadratic unbiased estimation (MINQUE) estimates. The MINQUE estimates are typically good estimates.

CTOLERANCE K

Specifies the convergence tolerance value for object function. K must be greater than 0. If this subsubcommand is not issued, Minitab uses the value 1E-7.

ETOLERANCE K

Specifies the convergence tolerance value for estimates. K must be greater than 0. If this subsubcommand is not issued, Minitab uses the value 1E-7.

LSPEC K

Specifies the lower specification limit. You must use LSPEC, USPEC, or both. If both LSPEC and USPEC are issued, USPEC must be greater than LSPEC. K must be numeric.

USPEC K

Specifies the upper specification limit. You must use LSPEC, USPEC, or both. If both LSPEC and USPEC are issued, USPEC must be greater than LSPEC. K must be numeric.

BPERCENT K

Specifies using K percent of the response above the lower spec or below the upper spec to estimate shelflife. K must be between 0 and 100. Minitab uses 50 percent if BPERCENT is not issued.

PVALUE K

Specifies the significance level that is used for model selection. K must be between 0 and 1. Minitab uses 0.25 if PVALUE is not issued.

ONESPEC

Specifies to use a one-sided significance level when you specify both specification limits. Use this subcommand when the response variable has two specification limits, but only one limit is relevant for the shelf life. For example, if the confidence level is 95% and you select this option, then the confidence level for the single bound that is relevant to the shelf life is also 95%. If you do not select this option, then the confidence level for both bounds is 97.5%. The single 95% bound gives a longer shelf life.

For example, a medication has an upper specification limit of 12.5 micrograms and a lower specification limit of 12 micrograms. In storage, the medication degrades, but never increases in strength. Only the lower specification limit is relevant to the shelf life.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

Box-Cox

BOXCOX [K]

Performs a Box-Cox transformation with a selected lambda. K is the value of lambda and must be between −5 and +5. If K is not given and you have a fixed batch, Minitab will find the optimal lambda. By default, Minitab rounds the optimal value.

Minitab cannot calculate the optimal lambda when batch is a random factor. Consequently, you must specify a lambda value for BOXCOX if you use RANDOM.

Graphs

GSHELFLIFE

Displays the shelf life plot. Use the subsubcommands to control how the batches are displayed.

INDIVIDUAL

Displays a graph for each individual batch.

ALL [K]

Displays all batches on one graph. If you specify K, each graph displays up to K batches.

RTYPE K

Specifies the type of residual to plot with the graph subcommands in K. If MARGINAL is issued, K must be 1 or 2.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

MARGINAL

Displays marginal residuals on the residual plots. If MARGINAL is not issued, the residual plots display conditional residuals for random batches.

GHISTOGRAM

Displays a histogram or individual value plot of the residuals, depending on the sample size.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

GFOURPACK
Displays a layout that includes a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs the order of the data.

Results

Displays a variety of tables in the output. The availability of some tables depends on whether BATCH is a fixed factor or a random factor.

NODEFAULT
Specifies that no default tables and graphs will be displayed.

TMETHOD
Displays the method table.

TFACTOR
Displays the name, number of levels, and the values for the batch factor in your model.

TITERATION
Displays the estimation iteration history table when batch is a random factor.

TVARIANCE
Displays the variance component table when batch is a random factor.

TCOVARIANCE
Displays the asymptotic variance covariance matrix of variance estimates when batch is a random factor.

TMSELECTION
Displays the model selection results, including the terms that are removed from the final model.

TANOVA
Displays the ANOVA table for a fixed batch.

TSUMMARY
Displays the summary of model table.

TCOEFFICIENTS
Displays the table of coefficients. Use FULL to display the full set of coefficients for the batch factor. Use NOFULL to show only the linearly independent coefficients. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL
Specifies to display the full set of coefficients for the batch factor

NOFULL
Specifies to not display the full set of coefficients for the batch factor

TRANDOM
Displays the random effect prediction table when batch is a random factor.

TEQUATION
Displays the regression equation table, or the marginal fitted equation table if batch is a random factor.

TDIAGNOSTICS [K]
Displays a table of fits and diagnostics. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations. If you do not specify K, Minitab uses the preference set in **File** > **Options** > **Linear Models** > **Display of Results**.

TCONDITIONAL [K]
Displays the conditional fits and residuals when batch is a random factor. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations. If you do not specify K, Minitab uses K = 0.

TSHELFLIFE
Displays the shelf life table.

TSIMPLE
Displays the simple versions of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

TEXPAND
Displays the expanded version of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

Storage

Use to store analysis results in the specified column or matrix. The availability and function of the storage commands depends on whether BATCH is a fixed factor or a random factor.

RESIDUALS C
Stores the residuals (fitted values – observed values) if batch is a fixed factor. Marginal residuals (marginal fitted values – observed values) if batch is a random factor.

SRESIDUALS C
Stores the standardized residuals if batch is a fixed factor. Standardized marginal residuals if batch is a random factor.

TRESIDUALS C
Stores the deleted Studentized residuals of the fit if batch is a fixed factor. If batch is a random factor, stores the studentized marginal residuals.

CRESIDUALS C
Stores the conditional residuals when batch is a random factor.

CSRESIDUALS C
Stores the conditional standardized residuals when batch is a random factor.

FITS C
Stores the fitted values if batch is a fixed factor. Marginal fitted values if batch is a random factor.

CFITS C
Stores the conditional fitted values when batch is a random factor.

HI C
Stores the leverages.

COOK C
Stores Cook's distance.

DFITS C
Stores the DFITS.

COEFFICIENTS C
Stores the estimated coefficients.

BLUP C
Stores the best linear unbiased predictor when batch is a random factor.

XMATRIX *M*
Stores the design matrix for fixed effects terms.

ZMATRIX *M*
Stores the design matrix for random effects terms.

COVARIANCE *M*
Stores the variance-covariance matrix of the variance component estimates if batch is a random factor.

FCOVARIANCE *M*
Stores the covariance matrix of fixed effect estimates.

NLINEAR: Session command for performing nonlinear regression

NLINEAR
Fits regression models in which the expected value of the response is a nonlinear function.

Use nonlinear regression to mathematically describe the nonlinear relationship between a response variable and one or more predictor variables. Specifically, use nonlinear regression instead of ordinary least squares regression when you cannot adequately model the relationship with linear parameters. Parameters are linear when each term in the model is additive and contains only one parameter that multiplies the term. Use this procedure for fitting models that are nonlinear in the parameters, storing regression statistics, examining residual diagnostics, generating point estimates, and generating prediction and confidence intervals.

Required subcommands

RESPONSE C
Specifies the response column. The column cannot be a text column.

CONTINUOUS C...C
Specifies one or more predictor columns. The columns cannot be text columns, and they must have the same number of rows as the response column.

PARAMETER *"name"* K...K

Use PARAMETER for each parameter in the model. The parameter name can match the name of a variable, but it cannot match the name of the response column or any predictor columns.

The parameter name must be followed by one or more starting values for the parameter. Minitab does not allow missing values.

If you give a parameter more than one starting value, Minitab evaluates the residual sum of squares (SSE) at all combinations, and uses the combination with the smallest value of the SSE.

EXPECTATION *expression*

Specifies the expectation function. The expression must use all of the parameters and predictors. All columns you use in the expression must be predictor columns.

If a parameter name looks like a variable name, Minitab interprets it as a variable name. To have Minitab interpret it as a name, you must put quotes around it. For example, a parameter you named k1 must be referred to as 'k1'.

Options

LOCK *parameters*

Locks one or more parameters at their starting values. If you give a locked parameter more than one starting value, the first value is used. You must leave at least one parameter unlocked. For example, LOCK "b1", locks the parameter b1.

WEIGHT C

Specifies weights for weighted regression in C. Weights must be nonnegative. Rows with zero or missing weight are omitted. Note that there is no provision for these weights to change from one iteration to the next.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

CONSTRAINT *constraint*

Specifies a constraint on a parameter. Repeat CONSTRAINT to specify multiple constraints. Each parameter can appear one time on one CONSTRAINT subcommand. For example, `CONSTRAINT 1200 < 'b1' < 1400` constrains the fitted value of parameter b1 to between 1200 and 1400.

MARQUARDT

Specifies the Levenberg-Marquardt algorithm rather than the default Gauss-Newton algorithm. Use this subcommand if the Gauss-Newton algorithm fails to converge.

TOLERANCE K [K]

K is the tolerance level for convergence, and it must be a positive number. The default value is 10^{-5}.

The convergence criterion is the relative offset orthogonality convergence criterion described in Bates and Watts (1981).

ITERATIONS K

Specifies the maximum number of iterations used in the optimization algorithm. K must be a positive integer. If you do not specify ITERATIONS, the default value is 200.

Predict

PCONTINUOUS *E..E*

PCONTINUOUS also computes the standard errors of the fitted values, 95% confidence intervals, and 95% prediction intervals. The number of arguments must equal the number of predictors in the model. If you include columns, they must have the same number of rows. Minitab assigns predictors in the same order as they appear on the CONTINUOUS subcommand.

If you omit PCONTINUOUS and include any of PFITS, PSEFITS, CLIMITS, or PLIMITS, then those subcommands are ignored.

The prediction interval computed by PREDICT assumes a weight of 1. If you used the WEIGHT subcommand with values other than 1, you should adjust the prediction interval values manually.

PFITS C

Stores the predicted values in C.

PSEFITS C

Stores the standard errors of the predicted values in C.

CLIMITS C C

Stores the upper and lower limits of the confidence intervals in C C. If you specify -1 or 1 with ITYPE then only one argument is allowed.

PLIMITS C C

Stores the upper and lower limits of the prediction intervals in C C. If you specify −1 or 1 with ITYPE then only one argument is allowed.

Tables

TMETHOD

Displays a table of information about the method and options. Includes the type of algorithm used, the maximum number of iterations, and the tolerance level for convergence.

TSTARTING

Displays the parameters' starting values.

TCONSTRAINTS

Displays a table of constraints for the parameters.

TITERATIONS

Displays the parameter estimates, and the residual sum of squares (SSE) at each iteration.

TEQUATION

Displays the regression equation.

TPARAMETERS
Displays the parameter estimates and their approximate standard errors.

CI
Adds a confidence interval for each parameter in the parameter estimates table.

TCORRELATION
Displays the correlation matrix for the parameter estimates.

TSUMMARY
Displays a test for lack of fit, and a summary of the fits that includes the number of iterations, SSE, the degrees of freedom for error, MSE, and S.

TPREDICTIONS
Displays a table of fitted values, confidence intervals, and prediction intervals for the new observations specified by PCONTINUOUS.

NODEFAULT
Suppresses the display of tables and graphs other than those specifically requested.

Graphs

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

GFCURVE
Displays a fitted curve plot. Minitab ignores the subcommand if there is more than one predictor.

GFCI
Overlays the confidence interval bands.

GFPI
Overlays the prediction interval bands.

Storage

RESIDUALS C
Stores the residuals (fitted values – observed values).

FITS C
Stores the fitted values, often called the Y-hats ($\hat{Y}$).

SPARAMETERS C
Stores the estimated parameters.

SSEPARAMETERS C
Stores the standard error of each parameter.

SGRID C...C
Stores each starting value combination and the corresponding initial sum of the squared residuals (SSE). Minitab performs the nonlinear regression analysis using the starting value combination that produces the smallest initial SSE.

OREG: Session command for performing orthogonal regression

OREG
Performs orthogonal regression in which both the response and predictor contain measurement error.

OREG uses orthogonal regression to fit a model to one predictor. Optionally, you can do the following:

- Store the residuals, fitted values, and variance estimates
- Generate predicted values for new observations
- Control display of results
- Generate several plots for residual analysis. For more information, go to Residual analysis and regression diagnostics on page 1261.

VARRATIO C
Specifies the error variance ratio (Y/X). K must be greater than 0.

Graphs

RTYPE K
Specifies the type of residual to plot with the graph subcommands in K.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

GFLINE
Displays a fitted line plot of the original data.

LSFIT
Displays the least squares fitted line on the fitted line plot. When you use LSFIT, both the orthogonal and the least squares regression equations are displayed on the graph.

Options

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PREDICT K

PREDICT C
Specifies the value of the predictor X. You can enter a constant or a column of values.

SPREDICT C
Stores the predicted values in the specified column.

SSTDEV C
Stores the standard deviations of the predicted values in the specified column.

Results

REVARRATIO
Displays the error variance ratio.

REQUATION
Displays the orthogonal regression equation.

RCOEFFICIENT
Displays the coefficient table.

RVARCOMP
Displays the error variance estimates.

RFITRESID
Displays the fitted values for the true predictor and true response as well as the residuals.

Storage

SRESIDUALS C
Stores the residuals in the specified column.

SSRESIDUALS C
Stores the standardized residuals in the specified column.

SCOEFFICIENT C
Stores the intercept and slope in the specified column.

SVARIANCE C
Stores the error variances for the predictor and response in the specified column.

SXFIT C
Stores the fitted values of the predictor.

SYFIT C
Stores the fitted values of the response.

PLS: Session command for performing partial least squares regression

PLS C...C = *termlist*
Use partial least squares regression (PLS) to relate a set of predictors to one or more response variables. Use PLS with ill-conditioned data, when predictors are highly collinear or predictors outnumber observations. Predictors can be either continuous or categorical. Use NCOMPONENTS to specify the number of components to use in the model. Use XVALIDATION to select the number of components that maximize the model's predictive ability. Use PCONTINUOUS and PCATEGORICAL to calculate fitted values for new observations using the PLS model.

CONTINUOUS C...C
Specifies which columns of variables are continuous. Columns must contain numerical or date/time data.

CATEGORICAL C...C
Specifies which, if any, columns of variables are categorical. Columns may be numerical, text, or date/time.

NCOMP K
Sets the maximum number of components to extract or cross-validate, which cannot exceed the number of predictors in your model or the number of observations minus one. Minitab extracts 10 or cross-validates

10 components by default, unless the number of predictors or the number of observations minus 1 is less than 10.

Options

XVALIDATION K

XVALIDATION C

XVALIDATION estimates your model's predictive ability by recalculating your model, leaving 1 or more observations out each time.

Use K to specify the number of observations to leave out each time the model is recalculated. Specifying 1 means that the model is recalculated as many times as there are observations.

Use C as a group identifier to indicate which observations are deleted together each time the model is recalculated. Observations with identical numbers in C are removed at the same time. The column must contain positive integers and equal the length of the predictor and response columns.

CODING K

CODING specifies the type of coding to use for categorical predictors.

- K = −1 for (−1, 0, +1) coding
- K = 1 for (1, 0) coding

REFERENCE C K...C K

REFERENCE changes the default coding for the categorical predictor columns. To change the default reference level, specify the categorical predictor column followed by the reference level. (You must enclose text and date/time levels in double quotes.) You can assign a reference level only when you use 1, 0 coding.

Prediction

Computes the fitted Ys, or $\hat{Y}$ for new observations using your PLS model.

PCONTINUOUS *E...E*

Use PCONTINUOUS for new continuous predictors.

PCONTINUOUS displays a table that contains the fitted Y's, standard errors of the fitted Y's, a 95% confidence interval, and a 95% prediction interval. E...E may be a list of constants or a list of columns, or a mixture of constants and columns. The number of predictors must equal the number of predictors in your PLS model.

PCATEGORICAL *E...E*

Use PCATEGORICAL for new categorical predictors.

PCATEGORICAL displays a table that contains the fitted Y's, standard errors of the fitted Y's, a 95% confidence interval, and a 95% prediction interval. E...E may be a list of constants or a list of columns, or a mixture of constants and columns. The number of predictors must equal the number of predictors in your PLS model.

RESPONSE C...C

Specifies response values for your new observations. Minitab calculates a test R^2 that indicates how well the model fits new observations.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PFITS C...C
Stores the fits from PCONTINUOUS and PCATEGORICAL.

PSEFITS C...C
Stores the estimated standard errors of the fits from PCONTINUOUS and PCATEGORICAL.

CLIMITS C, C...C, C
Stores the lower and upper confidence limits from PCONTINUOUS and PCATEGORICAL.

PLIMITS C, C...C, C
Stores the lower and upper prediction limits from PCONTINUOUS and PCATEGORICAL.

Storage of fits and residuals

FITS C...C
Stores the fitted values, often called the Y-hats ($\hat{Y}$).

CVFITS C...C
Stores the cross-validated fitted values.

RESIDUALS C...C
Stores the residuals (fitted values – observed values).

PRESIDUALS C...C
Stores the cross-validated residuals.

SRESIDUALS C...C
Stores the standardized residuals.

Storage of model information

COEFFICIENTS C...C
Stores the estimated coefficients, one stored column for each response.

SCOEFFICIENTS C...C
Stores the standardized coefficients, one stored column for each response.

HI C
Stores the severage values.

XDISTANCE C
Stores the distances from the x-model.

YDISTANCE C
Stores the distances from the y-model.

Storage of component information

XSCORES *M*
Stores the X-scores.

YSCORES *M*
Stores the Y-scores.

XLOADINGS *M*
Stores the X-loadings.

YLOADINGS *M*
Stores the Y-loadings.

XWEIGHTS *M*
Stores the X-weights.

XRESIDUALS *M*
Stores the X-residuals.

YRESIDUALS C...C
Stores the Y-residuals, one stored column for each response.

XCALCULATED *M*
Stores the X-calculated values.

YCALCULATED C...C
Stores the Y-calculated values, one stored column for each response.

Graphs for model evaluation

GSELECTIONPLOT
Displays a scatterplot of R^2 and predicted R^2 values vs the number of components

GFIT
Displays a scatterplot of fitted and cross-validated responses vs observed responses

GFCOEFFICIENT
Displays a projected scatterplot of the coefficients

GCCOEFFICIENT
Displays a projected scatterplot of the standardized coefficients

GDISTANCE
Displays a scatterplot of each observation's distance from the x-model and the y-model

Graphs for residual analysis

GHISTOGRAM
Displays a histogram of the standardized residuals

GNORMAL
Displays a normal probability plot of the standardized residuals

GRESIDUALS
Displays a scatterplot of residuals vs fitted values

GLEVERAGE
Displays a scatterplot of standardized residuals vs leverages

GFOURPACK
Displays a layout of a histogram of the standardized residuals, a normal probability plot of the standardized residuals, standardized residuals vs fitted values, and standardized residuals vs order of the data.

Graphs for model evaluation

GSCORE
Displays a plot of the x-scores of the first component versus the x-scores of the second component

GTHREEDPLOT
Displays a three-dimensional plot of the x-scores of the first component, second and third components

GLOADING
Displays a plot of the x-loadings of the first component vs the x-loadings of the second component

GXRESIDUALS
Displays a line plot of the x-residual matrix

GXCALCULATED
Displays a line plot of the x-calculated matrix.

Results

TMETHOD
Displays the number of components and analysis of variance table.

TANOVA
Displays the analysis of variance table.

TSELECTION
Displays the model selection and validation table.

TCOEFFICIENTS
Displays the unstandardized and standardized coefficients.

TFITS
Displays the table of fitted values and residuals.

TLEVERAGES
Displays the leverages and distances from the x- and y-model.

TPREDICTION
Displays the prediction table.

TXSCORES
Displays the x-score matrix.

TYSCORES
Displays the y-score matrix.

TXLOADINGS
Displays the x-loading matrix.

TYLOADINGS
Displays the y-loading matrix.

TXRESIDUALS
Displays the x-residual matrix.

TYRESIDUALS
Displays the y-residual matrix.

TXCALCULATED
Displays the x-calculated matrix.

TYCALCULATED
Displays the y-calculated matrix.

GZLM: Session command for fitting a binary logistic model or a Poisson model

GZLM

Creates a generalized linear model with the distribution model that you specify: logistic or Poisson. GZLM fits a model with one or more predictors using an iterative-reweighted least squares algorithm to obtain maximum likelihood estimates of the parameters. Use the logistic model for a binary response variable. Use the Poisson model for a Poisson response variable.

The predictors can be either factors (categorical predictors) or covariates (continuous predictors). Factors can be crossed. Covariates can be crossed with each other or with factors. For more information, go to How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC on page 1248. You can use stepwise methods to explore different terms in the model. You can use a test data set or k-fold cross-validation to validate the results.

Factor columns can be numeric or text. For a numeric factor, Minitab designates the lowest numeric value as the reference level. For a text factor, Minitab determines the reference level alphabetically. For an example, go to Entering data for factor variables on page 1225.

When the response variable column is binary, the highest value is the reference event. When the response variable column is text, reverse alphabetical order determines the event. For example, if you enter the response variable as success and failure, success is the event. For examples and information on entering the response data, go to Entering data for response variables on page 1225.

This analysis provides diagnostic plots, goodness-of-fit tests, and other diagnostic measures so you can assess the validity of your model. For more information, go to Generalized linear model diagnostics and residual analysis on page 1260.

Minitab excludes all observations with missing values on either the response variable or any of the predictors from all calculations.

BINOMIAL

Analyzes the relationship between a binary response variable that you specify in RESPONSE and the predictors that you with CONTINUOUS and CATEGORICAL. You specify the model with TERMLIST.

LOGIT

Use the logit link function.

NORMIT

Use the normit link function.

GOMPIT

Use the gompit link function. The gompit link function is also called the complementary log-log link function.

POISSON

Analyzes the relationship between a Poisson response variable that you specify in RESPONSE and the predictors that you with CONTINUOUS and CATEGORICAL. You specify the model with TERMLIST.

LOG

Use the natural log link function.

SQRT

Use the square root link function.

IDENTITY

Use the identity link function. The identity link function does not transform the response.

RESPONSE C [C]

Enters the response data as counts. With BINOMIAL, you can enter a column that contains the number of events first and then a column that contains the number of trials. For examples, go to Entering data for response variables on page 1225.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric or date/time and must match the length of the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric, text, or date/time and must match the length of the response column.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors may be repeated. Nested terms are not allowed. The model can be nonhierarchical. For examples, go to How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC on page 1248.

Options

FREQUENCY C

Indicates how often the data in a row occur. For examples, go to Entering data for response variables on page 1225.

CONSTANT

Fits the model with the intercept.

When you use CONSTANT, Minitab includes the intercept term in the model. For example, for a binomial logistic regression, Minitab fits the model

$g(p) = \beta_0 + \mathbf{x'\beta}$

If you do not specify CONSTANT or NOCONSTANT, then Minitab includes the intercept.

NOCONSTANT

Fits the model without the intercept.
When you use NOCONSTANT, Minitab omits the intercept term.

$g(p) = \mathbf{x'\beta}$

If you do not specify CONSTANT or NOCONSTANT, then Minitab includes the intercept.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

WALD

Use the Wald test for the ANOVA table.

LRT

Use the Likelihood ratio test for the ANOVA table.

ADJDEVIANCE

Use adjusted deviances for the tests of significance in the ANOVA table when the Likelihood ratio test is used.

SEQDEVIANCE

Use sequential deviances for the tests of significance in the ANOVA table when the Likelihood ratio test is used. The order of the terms in TERMLIST affects the statistical significance. The default is to use adjusted deviances.

PEARSON

Specifies the use of Pearson residuals. The default is to use deviance residuals.

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X'\ W\ X})^{-1}\ (\mathbf{X'\ W\ Y})$

EFFECT

Specifies an effect coding scheme for categorical predictors (+1, 0, –1). With effect coding, the coefficients compare the levels of the categorical variable to the mean. The default is binary coding (1, 0). With binary coding, the coefficients compare the levels of the categorical variable to a reference level.

HGROUP K

Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test with K. The default is 10.

REFERENCE C K...C K

Specifies the reference level for categorical predictor C with K. The coefficients compare the other levels of categorical variables to the reference level when you use binary coding.

LOCATION [K...K]

Standardize continuous predictors by subtracting the means. Optionally, you can subtract different constants that you specify with K...K. The order of the constants matches the order of the columns after CONTINUOUS.

SCALE [K...K]

Standardize continuous predictors by dividing by the standard deviations. Optionally, you can divide by different constants that you specify with K...K. The order of the constants matches the order of the columns after CONTINUOUS.

LEVELS K K...K K

Standardize continuous predictors by coding the low and high levels K K of each predictor to −1 1. The order of the constants matches the order of the columns after CONTINUOUS.

UNSTANDARDIZED

Use the original units for the continuous predictors. If you do not specify LOCATION, SCALE, LEVELS, or UNSTANDARDIZED, then Minitab uses the preferences that are set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

REVENT K

Specifies which is the response event with K. The value of K matches one of the values in the response column. Enclose text values in quotation marks. For example, to model the probability of a "Fail" in the response column, enter REVENT "Fail".

EVNAME K

When the RESPONSE command has two columns, you can specify the name of the event with K. Enclose text values in quotation marks.

ITERATION K

Specifies the maximum number of iterations for the optimization algorithm that estimates the coefficients. The default is 50.

TOLERANCE K

Specifies the tolerance level to keep a predictor in the model that is either highly correlated with another predictor or nearly constant. By default, K = 4 * 2.22E–012. K must be a positive number. Lowering tolerance by giving a very small argument can prevent Minitab from eliminating problematic predictor columns from the model.

Correlated predictors

1 – R-squared, where R-squared is the value resulting from regressing one predictor on the remaining predictors. If (1 – R-squared) is $\leq$ K, Minitab removes the predictor from the equation.

Constant predictors

The parameter for forcing in a variable that is nearly constant. If the coefficient of variation of a predictor (the standard deviation / the mean) is $\leq$ square root K, Minitab considers the predictor to be essentially constant and removes it from the equation.

CTOLERANCE K

Specifies the convergence criterion for the algorithm that estimates the coefficients. By default, K = 1E-8.

START C

Specifies a column that contains starting values for the coefficients in the optimization algorithm. The column begins with the estimate of the constant, then follows the order of the model after TERMS.

Stepwise selection

STEPWISE, FORWARD, BACKWARD, FINFORMATION and FVALIDATION perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

FVALIDATION

Specifies the use of forward selection with the test deviance R^2 as the criterion for the selection of the final model. When you use FVALIDATION, you must also use TEST. FVALIDATION is not compatible with FREQUENCY or ATEND.

When you use TEST, the procedure is similar to forward selection for other methods. At the end of each step, Minitab Statistical Software calculates the test R^2 statistic. At the end of the forward selection procedure, the model with the greatest test R^2 value is the final model. FVALIDATION stops under the same conditions as FINFORMATION.

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

GRSQUARE

After FVALIDATION, displays a graph of the R^2 statistic and the test deviance R^2 statistic for each step in the model selection procedure.

Validation

TEST C K

Specifies to perform validation with a test data set that you specify. The column must have two distinct values that indicate which rows are in the test data set and which rows are in the training data set. K is the value for rows that are in the test data set.

TESTK [K]

Specifies to perform validation with a randomly-selected test data set. The first K is the fraction of the data in the test data set, a value between 0 and 1. The second K is optional. The second K is a base for the random number generator. If you specify the base, you can get the same test data set again by specifying the same base.

CVTEST C

Specifies to perform k-fold cross validation with a test data set that you specify. Rows with the same value in the column are in the same fold.

CVTEST K [K]

Specifies to perform k-fold cross validation with randomly-selected folds. The constant is the number of folds. The optional constant is a base for the random number generator. If you specify the base, you can get the same folds again by specifying the same base.

Tables for all models

TMETHOD

Displays a table of information about the method and options.

TMSDETAILS

Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model.

If you do not specify FULL or NOFULL, then Minitab uses the preferences that are set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL

Specifies to display the coefficients, p-values, and model summary statistics for each step of the procedure. If you use the TEST subcommand, the model summary statistics include test R^2.

NOFULL

Hides these statistics.

TEQUATION

Displays the model equation.

SINGLE

SINGLE specifies to show one equation instead of a separate equation for each combination of the levels of the categorical predictors. The equation has a coefficient for each level of the categorical variables.

SEPARATE

Specifies to show separate equations for each combination of the levels of the categorical predictors. Each equation has a different constant term.

TDEVIANCE

Displays the ANOVA table for the Wald test or the likelihood ratio tests.

TSUMMARY

Displays the statistics that evaluate model fit, including deviance R^2.

TCOEFFICIENTS

Displays the coefficients and the p-values for the Wald normal approximation tests.

If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Specifies to show the coefficients for the reference level when you use (0, 1) coding for categorical predictors.

NOFULL

Hides these statistics.

TGOODNESS

Displays the Pearson and Deviance goodness-of-fit tests. For the binomial distribution models, also shows the Hosmer-Lemeshow goodness-of-fit test.

TDIAGNOSTICS [K]

Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

TSTEP

Displays the deviance at each iteration of the coefficients.

TEXPAND

Displays the expanded versions of the tables that have them. The expanded Analysis of Variance table adds the sequential deviances and percent contributions for each term. The Coefficients table adds confidence intervals for the coefficients. The Fits and Diagnostics table adds the standard error of the fit, the 95% confidence interval for the fit, the deleted residual, the HI, the Cook's D, and the DFITS.

TSIMPLE

Displays the simple versions of the Analysis of Variance table, the Coefficients table, and the Fits and Diagnostics table.

NODEFAULT

Hides default tables and graphs that you do not include the subcommand to display.

Tables for binomial distribution models

TRINFO

Displays the number of events, non-events, and the total. The table also indicates which response value is the reference event.

TODDS

Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model or for categorical predictors with (−1, 0,+1) coding.

INCREMENT K...K

Specifies the increment for the odds ratio for the predictor with K. For example, if the predictor is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram. Enter the values in the same order that the predictors follow CONTINUOUS. Enter 1 for a predictor to use the units of the data.

THOSMER

Displays the table of observed and expected values for the Hosmer-Lemeshow test.

TASSOCIATION

Displays the number of concordant and discordant pairs, the Somers' D, the Goodman-Kruskal Gamma, and Kendall's Tau-a.

Graphs for all models

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Displays residuals versus fitted values. For binary logistic regression, this plot is not available when the data are in response/frequency format because the pattern that appears is not a violation of the assumptions.

GORDER

Displays residuals versus the order of the worksheet. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK

Choose to display a layout of the residual plots. For Poisson models, the layout includes a histogram of residuals, a normal plot of residuals, a plot of residuals versus fits, and a plot of residuals versus order.

For the binary logistic model, the layout depends on the format of the data. If the data are in response/frequency format, the layout has a histogram of residuals, a normal plot of residuals, and a plot of residuals versus order. If the data are in event/trial format, the layout also has a plot of residuals versus fits.

GVARIABLE C...C

Displays the residuals versus the variables in C. Typically, you enter the columns for the predictors.

Graphs for binomial distribution models

GROC

Displays the receiver operating characteristic (ROC) curve for a binomial logistic regression. A footnote on the plot gives the area under the ROC curve. You can use the area under the ROC curve to compare models. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR).

Storage for all models

When you use the TEST subcommand, Minitab Statistical Software stores the following statistics for the training data set and not the test data set:

Subcommand	Statistic
TRESIDUALS	Deleted residuals
COOK	Cook's distances
DFITS	DFITS
SADJUSTEDRESP	Adjusted response values
XPWXINVERSE	X'WX inverse matrix of the estimated coefficients

FITS C

Stores the event probabilities.

RESIDUALS C

Stores the Deviance residuals by default. Stores the Pearson residuals if you use the subcommand PEARSON.

SRESIDUALS C

Stores the standardized Deviance by default. Stores the standardized Pearson residuals if you use the subcommand PEARSON.

TRESIDUALS C

Stores the deleted Deviance residuals by default. Stores the deleted Pearson residuals if you use the subcommand PEARSON.

COOK C

Stores the Cook's distances. If you use a test data set, the storage is for the training data set only.

DFITS C

Stores the DFITS.

HI C

Stores the leverages.

SADJUSTEDRESP C
Stores the final adjusted response values.

SIWEIGHTS C
Stores the internal weights used to estimate the model parameters.

SECOEFFICIENTS C
Stores the standard errors of the estimated coefficients.

COEFFICIENTS C
Stores the estimated coefficients.

XMATRIX *M*
Stores the design matrix.

XPWXINVERSE *M*
Stores the X'WX inverse matrix of the estimated coefficients.

When you use the TEST subcommand or the CVTEST subcommand, Minitab Statistical Software stores the following statistics for the training data set but not for the validation data set.

DBETA C
Stores the change in the estimated regression coefficients (delta b) when you delete all observations with a particular factor/covariate pattern. Use the delta b values to detect observations with a strong influence on the coefficients. If you use a test data set, the storage is for the training data set only.

DSBETA C
Stores the change in the estimated regression coefficients when you delete all observations with a particular factor/covariate pattern based on the standardized Pearson residual. If you use a test data set, the storage is for the training data set only.

DCHISQUARE C
Stores the change in the chi-square statistic when you delete all observations with a particular factor/covariate pattern. Observations that are poorly fit by the model have high delta chi-square values. If you use a test data set, the storage is for the training data set only.

DDEVIANCE C
Stores the change in the deviance when you delete all observations with a particular factor/covariate pattern. If you use a test data set, the storage is for the training data set only.

Storage for validation

Use the following commands to store statistics for validation with a test data set. These commands are sub-sub commands to the TEST subcommand. As such, these commands must come after TEST.

Use the following commands to store statistics for validation with a test data set or validation with k-fold cross validation.

SSAMPLE C
When CVTEST is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When TEST is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.

- C is a binary text column that contains Training and Test as the two values.

CVFIT C
Stores the fitted values from each fold's turn as the test data set.

CVRESID C
Stores the residuals from each fold's turn as the test data set.

BFIT: Session command for creating a binary fitted line plot

BFIT
Performs logistic regression on a binary response variable. A binary variable has only two possible values, such as presence or absence. BFIT fits a model with one continuous predictor using an iterative-reweighted least squares algorithm to obtain maximum likelihood estimates of the parameters.

For examples and information on entering the response data, go to Entering data for response variables on page 1225.

BFIT provides diagnostic plots, goodness-of-fit tests, and other diagnostic measures so you can assess the validity of your model. For more information, go to Generalized linear model diagnostics and residual analysis on page 1260.

Note Minitab excludes all observations with missing values on either the response variable or any of the predictors from all calculations.

RESPONSE C [C]
Enters the response data as counts. For examples and information on entering the response data, go to Entering data for response variables on page 1225.

CONTINUOUS C
Indicates the continuous predictor is in C.

LOGIT
Use the logit link function.

NORMIT
Use the normit link function.

GOMPIT
Use the gompit link function. The gompit link function is also called the complementary log-log link function.

Options

FREQUENCY C
Indicates how often the data in a row occur.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

REVENT K

Specifies which is the response event with K. The value of K matches one of the values in the response column. Enclose text values in quotation marks. For example, to model the probability of a "Fail" in the response column, enter REVENT "Fail".

EVNAME K

When the RESPONSE command has two columns, you can specify the name of the event with K. Enclose text values in quotation marks.

Tables

TMETHOD

Displays the link function.

TRINFO

Displays the number of events, non-events, and the total.

TDEVIANCE

Displays the ANOVA table.

TSUMMARY

Displays the statistics that evaluate model fit, including deviance R^2.

TCOEFFICIENTS

Displays the coefficients and the p-values for the Wald normal approximation tests.

TODDS

Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model or for categorical predictors with (−1, 0, +1) coding.

INCREMENT K

Specifies the increment for the odds ratio for the predictor with K. For example, if the predictor is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram.

TEQUATION

Displays the model equation.

TGOODNESS

Displays the deviance, Pearson, and Hosmer-Lemeshow tests. These tests assess the overall fit of the model.

TDIAGNOSTICS K

Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

NODEFAULT
Hides default tables and graphs that you do not include the subcommand to display.

Graphs

GFLINE
Displays the fitted line plot. The plot also shows the equation for the probabilities.

PROPORTIONS
Shows the data on the plot.

DCONFIDENCE
Shows the confidence interval for the fitted line on the plot.

RTYPE K
Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Displays residuals versus fitted values. This plot is not available when the data are in response/frequency format because the pattern that appears is not a violation of the assumptions.

GORDER
Displays residuals versus the order of the worksheet. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of the residual plots. If the data are in response/frequency format, the layout has a histogram of residuals, a normal plot of residuals, and a plot of residuals versus order. If the data are in event/trial format, the layout also has a plot of residuals versus fits.

GVARIABLE C...C
Displays the residuals versus the variables. Typically, you enter the columns for the predictors.

Storage

FITS C
Stores the event probabilities.

RESIDUALS C
Stores the Deviance residuals.

SRESIDUALS C
Stores the standardized Deviance residuals.

TRESIDUALS C
Stores the deleted Deviance residuals.

COEFFICIENTS C
Stores the estimated coefficients.

OLOGISTIC: Session command for performing ordinal logistic regression

OLOGISTIC C = C...C
Performs logistic regression on an ordinal response variable. Ordinal variables are categorical variables that have 3 or more possible levels with a natural ordering, such as strongly disagree, disagree, neutral, agree, and strongly agree.

OLOGISTIC fits a model with one or more predictors using an iterative-reweighted least squares algorithm to obtain maximum likelihood estimates of the parameters. OLOGISTIC assumes parallel lines; therefore, one set of coefficients is associated with the predictors. When this assumption is not valid, use NLOGISTIC on page 217, which generates separate logit functions.

The predictors can be either factors (nominal variables) or covariates (continuous variables). Factors can be crossed or nested. Covariates can be crossed with each other or with factors, or nested within factors. The model can include up to 9 factors and 50 covariates. Unless you specify a predictor in the model as a factor, Minitab assumes the predictor is a covariate. Model continuous predictors as covariates and categorical predictors as factors. For more information, go to How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC on page 1248.

Factor columns can be numeric or text. For a numeric factor, Minitab designates the lowest numeric value as the reference level. For a text factor, Minitab determines the reference level alphabetically. For an example, go to Entering data for factor variables on page 1225. To change the default coding, use the REFERENCE subcommand.

When the response variable column is numeric, OLOGISTIC orders the levels from smallest to largest. When the response variable column is text, OLOGISTIC orders the levels alphabetically. For examples and information on entering the response data, go to Entering data for response variables on page 1225. Use the ORDER subcommand to specify the ordering of the response variable.

FREQUENCY C
Enters the response data as counts. For examples, go to Entering data for response variables on page 1225.

Options

LOGIT
Specifies the logit link function (default).

NORMIT
Specifies the normit link function.

GOMPIT

Specifies the gompit link function (also called the complementary log-log function).

FACTORS C...C

Specifies which of the predictors are factors. Minitab assumes all variables in the model are covariates unless specified to be factors. Model continuous predictors as covariates and categorical predictors as factors.

REFERENCE C K, ..., C K

Changes the default coding for factors. To change the default reference factor level, specify the factor column followed by the reference level. (Text and date/time levels must be enclosed in quotes.) For a discussion and examples of the default coding scheme, go to Entering data for factor variables on page 1225.

Use the ORDER subcommand to specify the ordering of the response variable.

ORDER C

ORDER K...K

Specifies the order of the response values from lowest to highest. Values can be stored in C or K.

ITERATION K

Changes the maximum number of iterations that Minitab performs to reach convergence. The default value is 20. Minitab's logistic regression commands obtain maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

START C

Specifies the initial values for model parameters. The column containing the initial values must have one row for each estimated coefficient in the model. For OLOGISTIC, the starting values for all the constants appear first, followed by starting values for the predictors in the model.

Note Each degree of freedom must be in one row.

CTOLERANCE K [K]

Changes the convergence criteria. Both Ks must be positive numbers.

K	How convergence is attained
First K	Convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than K. The default is .0001.
Second K	Convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than K. The default is .0001.

TOLERANCE K

Specifies the tolerance level to keep a predictor in the model that is either highly correlated with another predictor or nearly constant. By default, K = 4 * 2.22E–012. K must be a positive number. Lowering tolerance by giving a very small argument can prevent Minitab from eliminating problematic predictor columns from the model.

Correlated predictors

1 – R-squared, where R-squared is the value resulting from regressing one predictor on the remaining predictors. If (1 – R-squared) is $\leq$ K, Minitab removes the predictor from the equation.

Constant predictors

The parameter for forcing in a variable that is nearly constant. If the coefficient of variation of a predictor (the standard deviation / the mean) is $\leq$ square root K, Minitab considers the predictor to be essentially constant and removes it from the equation.

Results

BRIEF K

Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage and displays the following error messages, warnings, prompts, and notes; graphs; WRITE to the screen.
1	Minitab displays the response information, logistic regression table, log-likelihood, and test that all slopes = 0.
2 (default)	In addition to the BRIEF 1 output, displays the Pearson and deviance tests.
3	In addition to the BRIEF 2 output, displays factor level values and tests for terms with more than 1 degree of freedom.

STEP

Displays the log-likelihood at each iteration of the parameter estimation process.

Storage of the characteristics of the estimated equation

COEFFICIENTS C

Estimated coefficients in a column.

SECOEFFICIENTS C

Standard errors of the coefficients in a column.

XPWXINVERSE *M*

Variance-covariance matrix of the estimated coefficients, expressed as $(X' W X)^{-1}$.

LOGLIKELIHOOD K

Last log-likelihood value.

Storage of the event probabilities and aggregated data

EPROBABILITY C...C

Predicted event probabilities. Specify one column for each distinct value of the response.

CUMPROBABILITY C...C

Cumulative event probabilities. Specify one column for each distinct value of the response minus one.

NTRIALS C

Number of trials for each factor/covariate pattern.

NOCCUR C...C

Number of occurrences for the j^{th} factor/covariate pattern. Specify one column for each distinct value of the response.

NLOGISTIC: Session command for performing nominal logistic regression

NLOGISTIC C = C…C

Performs logistic regression on a nominal response variable using an iterative-reweighted least squares algorithm to obtain maximum likelihood estimates of the parameters. Nominal variables are categorical variables that have 3 or more possible levels but with no natural ordering. For example, the levels in a food-tasting study may include crunchy, mushy, and crispy.

OLOGISTIC on page 214 assumes parallel lines. Therefore, one set of coefficients is associated with the predictors. When this assumption is not valid, use NLOGISTIC, which generates separate logit functions.

The predictors can be either factors (nominal variables) or covariates (continuous variables). Factors can be crossed or nested. Covariates can be crossed with each other or with factors, or nested within factors. The model can include up to 9 factors and 50 covariates. Unless you specify a predictor in the model as a factor using the FACTORS subcommand, Minitab assumes the predictor is a covariate. Model continuous predictors as covariates and categorical predictors as factors. For more information, go to How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC on page 1248.

Factor columns can be numeric or text. For a numeric factor, Minitab designates the lowest numeric value as the reference level. For a text factor, Minitab determines the reference level alphabetically. For an example, go to Entering data for factor variables on page 1225.

When the response variable column is numeric, NLOGISTIC defines the highest value as the reference event. When the response variable column is text, NLOGISTIC defines the event by reverse alphabetical order. For example, if you enter the response variable as crunchy, mushy, and crispy, NLOGISTIC defines mushy as the event. For examples and information on entering the response data, go to Entering data for response variables on page 1225.

Use the REFERENCE subcommand to change the default coding for both factor and response variables.

NLOGISTIC provides goodness-of-fit tests and other diagnostic measures so you can assess the validity of your model.

FREQUENCY C

Enters the response data as counts. For examples and information on entering the response data, go to Entering data for response variables on page 1225.

Options

FACTORS C...C

Specifies which of the predictors are factors. Minitab assumes all variables in the model are covariates unless specified to be factors. Continuous predictors must be modeled as covariates; categorical predictors must be modeled as factors.

REFERENCE C K, ..., C K

Changes the default coding for factor and response columns. To change the default reference factor level, specify the factor column followed by the reference level. (Text and date/time levels must be enclosed in quotes.) For a discussion and examples of the default coding scheme, go to Entering data for factor variables on page 1225.

To change the default event, specify the response column followed by the event value.

START C

Specifies the initial values for model parameters. The column containing the initial values must have one row for each estimated coefficient in the model. For NLOGISTIC, the starting values for logit 1 appear before logit 2, and so on. For each logit, the starting value for the constant appears before the starting values for the predictors in the model.

Note Each degree of freedom must be in one row.

ITERATION K

Changes the maximum number of iterations that Minitab performs to reach convergence. The default value is 20. Minitab's logistic regression commands obtain maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

CTOLERANCE K K

Changes the convergence criteria. Both Ks must be positive numbers.

K	How convergence is attained
First K	Convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than K. The default is .0001.
Second K	Convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than K. The default is .0001.

TOLERANCE K

Specifies the tolerance level to keep a predictor in the model that is either highly correlated with another predictor or nearly constant. By default, K = 4 * 2.22E–012. K must be a positive number. Lowering tolerance by giving a very small argument can prevent Minitab from eliminating problematic predictor columns from the model.

Correlated predictors

1 – R-squared, where R-squared is the value resulting from regressing one predictor on the remaining predictors. If (1 – R-squared) is $\leq$ K, Minitab removes the predictor from the equation.

Constant predictors

The parameter for forcing in a variable that is nearly constant. If the coefficient of variation of a predictor (the standard deviation / the mean) is $\leq$ square root K, Minitab considers the predictor to be essentially constant and removes it from the equation.

Results

BRIEF K

Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage and displays the following error messages, warnings, prompts, and notes; graphs; WRITE to the screen.
1	Minitab displays the response information, logistic regression table, log-likelihood, and test that all slopes = 0.
2 (default)	In addition to the BRIEF 1 output, displays the Pearson and deviance tests.
3	In addition to the BRIEF 2 output, displays factor level values and tests for terms with more than 1 degree of freedom.

STEP
Displays the log-likelihood at each iteration of the parameter estimation process.

Storage of the characteristics of the estimated equation

COEFFICIENTS C
Estimated coefficients in a column.

SECOEFFICIENTS C
Standard errors of the coefficients in a column.

XPWXINVERSE *M*
Variance-covariance matrix of the estimated coefficients, expressed as $(X' W X)^{-1}$.

LOGLIKELIHOOD K
Last log-likelihood value.

Storage of the event probabilities and aggregated data

EPROBABILITY C...C
Predicted event probabilities. Specify one column for each distinct value of the response.

NTRIALS C
Number of trials for each factor/covariate pattern.

NOCCUR C...C
Number of occurrences for the j^{th} factor/covariate pattern. Specify one column for each distinct value of the response.

ANOVA

ONEWAY: Session command for performing a one-way ANOVA

ONEWAY

Performs a one-way analysis of variance. Your data can be arranged in one of two ways:

- Response data are in one column for all factor levels
- Response data are in a separate column for each factor level

Minitab performs the traditional one-way ANOVA procedure if you assume that the variance is constant across all groups. However, if your data exhibit unequal variances between groups, Minitab can perform Welch's ANOVA. To perform Welch's ANOVA, use the WELCH subcommand.

Use the comparison method subcommands to determine which group means differ and by how much.

The factor column may be numeric or text, and may contain any value. The levels do not need to be in any special order. SET is especially useful for entering data that follow a pattern, such as subscripts (factor level values). For examples, go to Entering patterned data for the SET session command on page 1226. (TSET and DSET can be used to enter text and date/time data, respectively, that follow a pattern.)

Variables and terms to include in the model

Specifies information about the variables and terms to include in the model. How you use these commands depends on how your data are arranged in the worksheet. Your response data can either be in one column for all factor levels or the response data can be in a separate column for each factor level.

RESPONSE C

RESPONSE C...C

Specifies the column or columns that contains the response data. The column or columns must be numeric or date/time. If your response data are in one column, specify the column with C. If your response data are in a separate column for each factor level, specify all of the response columns with C...C.

CATEGORICAL C

Required when your response data are in one column for all factor levels. For this data arrangement, you must use CATEGORICAL to specify the single column that contains the categorical factor. If you specify more than one column for RESPONSE, CATEGORICAL is not allowed.

Options

CONFIDENCE K

Specifies a confidence level for the interval plot and means table, but not the multiple comparisons, in K. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval for the interval plot and means table.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

WELCH

Specifies the Welch method. Use WELCH when you do not assume that all groups have equal variances. One-way ANOVA with equal variances is slightly more powerful than Welch's ANOVA, but serious error can result if the variances are not equal.

Comparisons

TUKEY [K]

Calculates all pairwise differences between level means using Tukey's method (also called Tukey's HSD or Tukey-Kramer method) with the family error rate specified as K. If K is not specified, Minitab uses the default family error rate of .05.

To obtain output for this comparison method, you must issue at least one of the following subcommands: TGROUPING, TMTEST, GMCI.

FISHER [K]

Calculates all pairwise differences between level means using Fisher's LSD procedure with the individual error rate specified as K. If K is not specified, Minitab uses the default individual error rate of .05.

To obtain output for this comparison method, you must issue at least one of the following subcommands: TGROUPING, TMTEST, GMCI.

DUNNETT [K] K

DUNNETT C

Calculates the difference between each treatment mean and a control mean. The family error rate is specified by the first K. How you specify the control group depends on how your data are arranged.

Specify the control group as follows:

- If your response data are in one column, specify the control group with the second K.
- If your response data are in a separate column for each factor level, specify the control group column with C.

If the error rate is not specified, Minitab uses the default family error rate of .05.

When the column of levels contains text data, the control level will also be text so remember to enclose the value you enter for K in double quotes.

To obtain output for this comparison method, you must issue at least one of the following subcommands: TGROUPING, TMTEST, GMCI.

MCB [K] K

Calculates the difference between each level mean and the best of the other level means. The family error rate is specified by the first K and the type of best is specified by the second K. If the error rate is not specified, Minitab uses the default family error rate of .05.

There are two choices for "best." If the smallest mean is considered the best, set K = -1; if the largest is considered the best, set K = 1.

To obtain output for this comparison method, you must issue at least one of the following subcommands: TGROUPING, TMTEST, GMCI.

GAMES [K]

Calculates all pairwise differences between level means using Games-Howell method with the family error rate specified as K. If K is not specified, Minitab uses the default family error rate of .05. GAMES is only available when WELCH is issued.

To obtain output for this comparison method, you must issue at least one of the following subcommands: TGROUPING, TMTEST, GMCI.

Results

NODEFAULT

Specifies that no default tables and graphs will be displayed.

TEXPAND

Displays the expanded versions of the ANOVA table, Model summary, and Multiple comparisons tables.

TSIMPLE

Displays the simple version of all tables.

TMETHOD

Displays the method table.

TFACTOR

Displays the name, number of levels, and the values for the categorical factor.

TANOVA

Displays the ANOVA table.

TSUMMARY

Displays the summary of model table.

TMEANS

Displays the table of group means.

TGROUPING

Displays the grouping information table for each comparison method subcommand that you issue. This table highlights the significant and non-significant comparisons for each comparison method.

TMTEST

Displays the multiple comparison test table for each comparison method subcommand that you issue. This table displays the hypothesis test form of the comparison output, which includes the differences of the means, the numeric values for the confidences intervals, and the adjusted p-values.

Graphs that display information by group

GINTPLOT

Displays an interval plot of the group means.

GINDPLOT
Displays an individual value plot for each group.

GBOXPLOT
Displays a boxplot for each sample.

Graphs for the residuals

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals versus fitted values, and residuals versus order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Other graphs

GMCI
Displays an interval plot of the differences between group means for each comparison method subcommand that you issue.

Storage

FITS C

FITS C...C
Stores fitted values. The fitted values are the level means. Specify one storage column C for each column that contains response data.

RESIDUALS C

RESIDUALS C...C
Stores residuals. Residual = (response – fit). Specify one storage column C for each column that contains response data.

ANOM: Session command for creating an analysis of means chart

ANOM C

Displays an analysis of means (ANOM) chart. You must choose one distribution that the response data follows: normal (one-way or two-way), binomial, or Poisson.

For ANOM with a normal distribution and one factor, a single plot displaying the means for each level of the factor, a center line which is the grand mean, and upper and lower decision limits is produced. If you use ANOM with a normal distribution and two factors, three plots are given: one showing the interaction effects, one showing the main effects for the first factor, and one showing the main effects for the second factor. If you choose an ANOM for a binomial distribution or a Poisson distribution, a single plot displaying the means for each level of the factor, a center line which is the grand mean, and upper and lower decision points is produced.

The subcommands NORMAL, BINOMIAL, and POISSON replace the commands %ANOM, %BANOM, and %PANOM. respectively.

NORMAL C [C]

For normal data, each row of the response column represents an observation for measurement data. Response data are typically measurement data, such as weight or moisture content.

BINOMIAL K

For binomial data, the values in the response column are the numbers of defectives found in each sample. These values must be positive integers (0 or greater). You can include up to 500 samples.

POISSON

For Poisson data, the values in the response column are the numbers of defects that are found in each sample. These values must be positive integers (0 or greater). You can include up to 500 samples.

ALPHA K

The decision lines on an ANOM chart are based on an experiment-wide error rate, similar to what you might use when making pairwise comparisons or contrasts in an ANOVA. By default, a rate of 0.05 is used. Use ALPHA to enter a number between 0 and 1. Values greater than or equal to 1.0 are interpreted as percentages.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

When you use TITLE with %graphs or the DENDROGRAM subcommand, you can use it only one time, and you cannot use any of the TITLE subcommands available with the graph commands.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

ANOVA: Session command for performing a balanced ANOVA

ANOVA C...C
Performs univariate and multivariate analysis of variance.

For one-way analysis of variance, you can have unbalanced designs. For multi-way analysis of variance, you must have balanced designs (all cells must have the same number of observations). The command GLM on page 228 analyzes balanced and unbalanced designs. However, if your design is balanced, ANOVA is faster and requires less space.

Factors can be crossed or nested, fixed or random. ANOVA calculates all exact F-tests, displays expected mean squares, and estimates variance components. You can specify your own tests, store residuals and fitted values, and display cell and marginal means. You can analyze up to 50 response variables and up to 9 factors on one ANOVA command.

Minitab verifies that your model is valid and displays a message if it is not. Minitab also verifies that your data set is balanced.

Note Balanced data are not required for one-factor models.

If you do not use the subcommand RANDOM, ANOVA fits a fixed effect model. In this case, the F-statistic for a term is always (MS term) / (MSE). However, when some of your factors are random, the denominator of an F-test, in general, is not MSE. For more information, see the subcommands RANDOM and EMS.

You can use the subcommands MANOVA, SSCP, EIGEN, PARTIAL, and NOUNIVARIATE to do multivariate analysis of variance. Hotelling's T-squared test can be performed as a special case of ANOVA with MANOVA.

ANOVA can analyze very complex designs, but, if you have a simple model, ANOVA is very easy to use.

Fore more information, go to How to specify the model in ANOVA on page 1252 and How to enter data for ANOVA and GLM on page 1241.

Model

RANDOM C...C

Specifies which factors are random. Do not include interaction terms or nested factors. If at least one factor in an interaction term is random, then Minitab considers the term to be random. Any term that is nested within a random factor is considered random.

RESTRICT

There are two mixed model analysis of variance models: one requires the crossed, mixed terms to sum to 0 over subscripts that correspond to fixed effects (Minitab refers to this as the restricted model.), and the other does not. Most textbooks and BMDP's program 8V use the restricted model. SAS uses the unrestricted model.

By default, ANOVA fits the unrestricted model. This subcommand instructs the program to fit the restricted model, which assumes that the mixed interaction terms are restricted to sum to 0 over the fixed effects. For more information, go to Restricted and unrestricted mixed models on page 1227.

Results

EMS

Displays a table that contains expected mean squares, estimated variance components, and the error term (the denominator) used in each exact F-test. If no exact F-test for a term exists, then use the expected mean squares to determine how to construct an approximate F-test using the subcommand TEST.

The estimates of the variance components are the usual unbiased analysis of variance estimates. They are obtained by setting each calculated MS equal to its EMS. This gives a system of linear equations in the unknown variance components. This system is then solved. Unfortunately, this method can result in negative estimates, which should be set to 0. However, Minitab displays the negative estimates because they sometimes indicate that the model that is being fit is inappropriate for the data.

Terms that are fixed do not have variance components estimated.

MEANS *termlist*

Displays a table of means corresponding to each term listed. For example, the following command language displays four tables, one table for each main effect, A, B, D, and one for the three-way interaction, A*B*D.

```
ANOVA Y = A|B|C(A B) |D;
 MEANS A B D A*B*D
```

Terms listed on the MEANS statement must also be terms listed in the model. If you specify more than one MEANS subcommand, then Minitab uses the last one. For more information, go to How to specify the model in ANOVA on page 1252.

TEST *termlist / errorterm*

Calculates univariate F-tests. Termlist is a list of terms in the model. Each term is used as the numerator in a test. Errorterm is a term in the model to be used as the denominator in all the test. Alternatively, errorterm can be a linear combination of terms in the model and MSE (denoted by the reserved word ERROR). Minitab calculates the F-ratios and p-values for you. Here are two examples.

```
TEST A/A*B + A*C + B*C - 2*A*B*C - 2*ERROR
TEST A B/A*B
```

Minitab constructs the synthetic denominator MS for the F-test. This denominator is the linear combination of the MSs that are specified on TEST, and should have the same expectation as the numerator MS under the hypothesis being tested. Suppose the denominator is the linear combination:

$$\sum a_j MS_j$$

Then it has approximate degrees of freedom calculated by

$$DF\ (denominator) = \frac{[\sum a_j MS_j]^2}{\sum [(a_j MS_j)^2 / DF_j]}$$

where DF_i is the degrees of freedom for the term MS_i. For more information, go to Restricted and unrestricted mixed models on page 1227.

Storage

FITS C...C

Stores fitted values, using one column for each response. If you specify a full model (you include all interaction terms), then the fitted values are just the cell means. If you fit a reduced model, the fitted values are not the cell means, but are the sum of the least squares estimates of the effects in the model.

RESIDUALS C...C

Stores residuals, using one column for each response variable. Residual = (response – fit).

Graphs

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C

Displays a separate graph for the residuals versus each specified column.

Balanced MANOVA

MANOVA *[termlist [/ errorterm]]*

Performs four multivariate tests—Wilks' test, Lawley-Hotelling test, Pillai's test, and Roy's largest root test—for each term in the model. If you include a termlist, MANOVA performs the four multivariate tests for each term listed. Hotelling's T-squared test can be performed as a special case of ANOVA with MANOVA.

If you specify an errorterm, then it must be a single term that is in the model. MANOVA then uses this errorterm in all tests. If you do not specify an errorterm, then Minitab determines an appropriate errorterm, as in the univariate case.

All four tests are based on two SSCP (sums of squares and cross products) matrices: H = the hypothesis matrix and E = the error matrix. There is one H associated with each term in termlist. E is the matrix associated with the error for the test.

SSCP

Displays the hypothesis matrix, H, that corresponds to each term that is specified by MANOVA, and the error matrix, E.

EIGEN

Displays a table that contains the eigenvalues and eigenvectors for the (nonsymmetric) matrix, E-1 H. These are the eigenvalues that are used to calculate the four MANOVA tests. A separate table is displayed for each term that is specified on MANOVA.

Note If an eigenvalue is repeated, then the corresponding eigenvectors are not unique. In this case, the eigenvectors that Minitab displays and those in books or other software might not agree. However, the MANOVA tests are always unique.

PARTIAL

Displays a matrix of partial correlations. These are the correlations among the residuals or, equivalently, the correlations among the responses conditioned on the model. The formula for the matrix is W**-.5 E W**-.5, where E is the error matrix and W has the diagonal of E as its diagonal and 0s elsewhere.

NOUNIVARIATE

Suppresses the univariate output. Only the multivariate output is displayed.

GLM: Session command for fitting the general linear model

GLM

Fits the general linear model (GLM). Using GLM, you can perform analysis of variance with balanced and unbalanced designs, analysis of covariance, and regression.

Factors can be crossed or nested, fixed or random. Covariates can be crossed with each other or with factors, or nested within factors. You can store residuals, fitted values, and many other diagnostics, and display cell and marginal means.

GLM does tests for fixed effects models automatically. The TEST subcommand allows you to calculate the appropriate univariate F-tests for mixed models.

Output contains both the sequential sums of squares and the adjusted sums of squares (that is, each term is fit after all other terms in the model). Automatic tests are done using the adjusted SS, assuming all factors are fixed. Observations that are considered unusual are displayed. You can also display coefficients, along with their standard deviations and t-values, and cell and marginal means.

Calculations are performed using a regression approach. First a "full rank" design matrix is formed from the factors and covariates. The columns of the design matrix are used as predictors. Then each response variable is regressed on these columns.

You can store the residuals, fitted values, prediction and confidence intervals, and many other diagnostics for further analysis.

For more information, go to How to specify the model for GLM on page 1250 and How to enter data for ANOVA and GLM on page 1241.

Options

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

CONTINUOUS C...C

Specifies the continuous predictors if you have any. The column or columns must be numeric or date/time and must match the length of the response column.

CATEGORICAL C...C

Specifies the categorical predictors if you have any. The column or columns can be numeric, text, or date/time and must match the length of the response column.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors may be repeated. Nested terms are not entered in the term list. The model can be nonhierarchical when there are no random factors.

RANDOM C...C

Specifies random factors. Random factors must also be specified in CATEGORICAL.

NESTED C (C...), ..., C (C...C)

Specifies the nesting relationships. Parentheses are used for nesting. For example, when B is nested within A, use B (A), and when C is nested within both A and B, use C (A, B). Terms in parentheses are always factors in the model and are listed with commas between them. Thus, D (A, B, E) is correct but D (A*B E) and D (A*B*E) are not. Also, parentheses are not used inside parentheses. Thus, C (A, B) is correct but C (A, B (A)) is not.

WEIGHT

Performs a weighted regression. Weights cannot be used with optimal Box-Cox transformation. Column must be numeric with nonnegative values. Length must match response column length.

An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{Y})$

This is equivalent to minimizing the weighted SS Error:

$\Sigma w_i (Y - \hat{Y})^2$ where w_i is the weight in row i

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

TOLER K

Specifies the tolerance level for collinearity and constant check. Use to force Minitab to keep a predictor in the model which is either highly correlated with another predictor or nearly constant. The default k = 4*2.22E-16.

SSQUARES

Specifies sequential sum of squares for tests in the ANOVA table. The default is the adjusted sums of squares.

MEANS *termlist*

Calculates least squares means for specified terms. Terms must be in the model.

TEST *termlist / error term*

Use TEST to specify your own tests. Termlist is a list of terms in the model. List them exactly as on the TERMS subcommand. Error term is the term to be used as the denominator for the F-test. This can be a term in the model, a linear combination of terms in the model, or MSE (denoted by the word ERROR). Only available via command line.

EFFECT

Specifies the effect coding (–1, 0, +1) scheme for categorical predictors. If you do not specify either EFFECT or BINARY, then Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

BINARY

Specifies the binary coding (1, 0) scheme for categorical predictors. If you do not specify either EFFECT or BINARY, then Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

REFERENCE C K...C K

Changes the default coding for the categorical predictor columns. To change the default reference factor level, specify the factor column followed by the reference level. (You must enclose text and date/time levels in double quotes.) You can assign a reference level only when you use the binary coding (1, 0) scheme.

Options for standardizing the continuous predictor

Use this set of subcommands to standardize the continuous predictors in your model. You can use SCALE and LOCATION in conjunction with each other. LEVELS is mutually exclusive with the LOCATION and SCALE subcommands. If you do not specify LOCATION, SCALE, LEVELS, or UNSTANDARDIZED, then Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

LOCATION [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by subtracting a constant from each predictor. If you do not specify any arguments, the mean of each predictor column is subtracted. K specifies to subtract a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

SCALE [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by dividing each predictor by a constant. If you do not specify any arguments, each predictor column is divided by the standard deviation. K specifies to divide by a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

LEVELS [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by DOE-type coding for the specified low and high levels K K...K K. The number of arguments must be twice the number of continuous predictors.

UNSTANDARDIZED

Specifies the analysis is to be performed on the original predictors.

Box-Cox

BOX-COX [K]

Performs a Box-Cox transformation with a specified lambda. K is the value of lambda and must be between −5 and +5. If K is not given, Minitab will find the optimal lambda. By default, Minitab rounds the optimal value.

Minitab cannot calculate the optimal lambda for stepwise regression or when the model contains random factors. Consequently, you must specify a lambda value for BOXCOX if you use RANDOM, STEPWISE, FORWARD, or BACKWARD.

Stepwise

`FINFORMATION`, `STEPWISE`, `FORWARD`, and `BACKWARD` perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value

of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS
Specifies that both categorical and continuous terms have to be hierarchical.

ATEND
Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS
Specifies that the model is hierarchical at every step.

SINGLE
Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE
Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

RTYPE K
Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM
Displays a histogram or individual value plot of the residuals, depending on the sample size.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Results

NODEFAULT
Specifies that no default tables and graphs will be displayed.

TMETHOD
Displays the method table.

TFACTOR
Displays the name, number of levels, and the values for all categorical factors in your model.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Displays the coefficients, p-values, Mallows' Cp, and model summary statistics for each step of the procedure.

NOFULL
Hides these statistics.

TEQUATION
Displays the regression equation table. Minitab will display up to 50 equations. If you do not specify SINGLE or SEPARATE, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

SINGLE
If you want to view a single equation, rather than a separate equation for each factor level combination, use SINGLE.

SEPARATE
To view the separate equations, use SEPARATE.

TCOEFFICIENTS
Displays the table of coefficients. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL
Displays the complete set of coefficients for categorical predictors.

NOFULL
Displays only the linearly independent coefficients.

TSUMMARY
Displays the summary of model table.

TANOVA
Displays the ANOVA table.

TDIAGNOSTICS K
Displays a table of diagnostics. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations.

TDW
Displays Durbin-Watson statistics.

TEXPAND
Displays the expanded version of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

TSIMPLE
Displays the simple version of the ANOVA table, table of coefficients, model summary table, and table of unusual observations.

TEMS
Displays the table of expected mean squares, estimated variance components, and error term in each F-test.

TMEANS
Displays the table of least squares means.

TVARIANCE
Displays the table of variance components.

TTEST
Displays the table of your own tests.

Storage for Box-Cox

BCRESP
Stores the Box-Cox transformation of the response in C.

BFITS
Stores the fits for the original response.

BSMEANS
Stores the means for the original response.

Storage of fits and residuals

RESIDUALS
Stores the residuals (fitted values – observed values).

SRESIDUALS
Stores the standardized residuals.

TRESIDUALS
Stores the deleted Studentized residuals.

HI
Stores the leverages.

COOK
Stores Cook's distance.

DFITS
Stores DFITS.

Storage of characteristics of the estimated equation

COEFFICIENTS
Stores the estimated coefficients.

FITS
Stores the fitted values, often called the Y-hats ($\hat{Y}$).

SMEANS
Stores the least square means for the terms specified by MEANS.

XMATRIX
Stores the design matrix for regression model.

REML: Session command for fitting a mixed effects model

REML

By default, fits a model with random factors with the Restricted Maximum Likelihood method (REML). Using REML, you can perform analysis of variance on random factors with balanced or unbalanced designs. With the subcommand MLE, you can use maximum likelihood estimation.

For REML, at least 1 factor must be random. Other factors can be fixed or random. Covariates can be crossed with each other, with factors, or nested within factors. You can store residuals, fitted values, and other diagnostic statistics.

REML tests for fixed effects with either the Kenward-Roger approximation or the Satterthwaite approximation. To use the Satterthwaite approximation, use the subcommand SATTERTHWAITE. To show the tests, use the subcommand TFIXEDEFFECT.

Tests use the adjusted sums of squares. Observations with standardized residuals that have absolute values greater than 2 are in a table of unusual observations. to show the table, use the subcommand TDIAGNOSTICS.

REML has 4 required subcommands:

- RESPONSE
- RANDOM
- CATEGORICAL
- TERMS

For more information, go to the descriptions of the individual subcommands.

Options

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

CATEGORICAL C...C

Specifies the categorical predictors if you have any. The column or columns can be numeric, text, or date/time and must match the length of the response column.

CONTINUOUS C...C

Specifies the continuous predictors if you have any. The column or columns must be numeric or date/time and must match the length of the response column.

RANDOM C...C

Specifies random factors. Random factors must also be specified in CATEGORICAL. At least one factor must be random.

NESTED C (C...), ..., C (C...C)

Specifies the nesting relationships. Parentheses are used for nesting. For example, when B is nested within A, use B (A), and when C is nested within both A and B, use C (A, B). Terms in parentheses are always factors in the model and are listed with commas between them. Thus, D (A, B, E) is correct but D (A*B E) and D (A*B*E) are not. Also, parentheses are not used inside parentheses. Thus, C (A, B) is correct but C (A, B (A)) is not.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors may be repeated. Nested terms are not entered in the term list. The model can be nonhierarchical.

WEIGHT C

Performs a weighted regression. Weights cannot be used with optimal Box-Cox transformation. Column must be numeric with nonnegative values. Length must match response column length.

An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{Y})$

This is equivalent to minimizing the weighted SS Error:

$\Sigma w_I (Y - \hat{Y})^2$ where w_I is the weight in row I

REML

Specifies to use the restricted maximum likelihood estimation method to estimate the variance components.

MLE

Specifies to use the maximum likelihood method to estimate the variance components. Usually, you use the REML method

MAXITER K

Specifies the maximum number of iterations for the Newton-Raphson algorithm to estimate the variance components.

STARTING C...C or K...K

Specifies the starting estimates for the variance components for each response in C. The order of the values in each column must be the same order of the random terms in TERMS. The last value in C is for the error term. If you have only one response or you want to use the same starting estimates for all responses, you can enter numbers K...K.

CTOLERANCE K
Specifies the convergence tolerance value in K for the objective function. The objective function is -2 log likelihood.

ETOLERANCE K
Specifies the convergence tolerance value in K for the estimates of the variance components.

KENWARDROGER
Use Kenward-Roger approximation to estimate the denominator degrees of freedom for the fixed effect tests.

SATTERTHWAITE
Use Satterthwaite approximation to estimate the denominator degrees of freedom for the fixed effects.

DROWSTAT
Computes statistics that require row deletion:

- PRESS
- R-Sq(Pred)
- Cook's D
- DFITS
- Marginal deleted residuals
- Marginal conditional residuals

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K
Specifies the type of confidence interval.

Value of K	Type of confidence interval
−1	Lower bound
0	Two-sided
1	Upper bound

TOLERANCE K [K]
Specifies the tolerance level for the collinearity between variables and whether a variable is constant. The first K specifies the tolerance for collinearity. The default value is 1E−18. The second K is for the tolerance for whether a variable is constant. The default value is 1E−21. Lower the tolerance to keep variables in the model that Minitab would exclude with the default values.

MEANS *termlist*
Calculates least squares means for specified terms. Terms must be in the model.

Options for standardizing the continuous predictor

Use this set of subcommands to standardize the continuous predictors in your model. You can use SCALE and LOCATION in conjunction with each other. LEVELS is mutually exclusive with the LOCATION and SCALE subcommands. If you do not specify LOCATION, SCALE, LEVELS, or UNSTANDARDIZED, then Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

LOCATION [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by subtracting a constant from each predictor. If you do not specify any arguments, the mean of each predictor column is subtracted. K specifies to subtract a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

SCALE [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by dividing each predictor by a constant. If you do not specify any arguments, each predictor column is divided by the standard deviation. K specifies to divide by a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

LEVELS [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by DOE-type coding for the specified low and high levels K K...K K. The number of arguments must be twice the number of continuous predictors.

UNSTANDARDIZED

Specifies the analysis is to be performed on the original predictors.

Graphs

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

CONDITIONAL

Specifies to plot residuals for conditional fits.

MARGINAL

Specifies to plot residuals for marginal fits.

GHISTOGRAM

Displays a histogram or individual value plot of the residuals, depending on the sample size.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GORDER

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Tables

NODEFAULT
Specifies that no default tables and graphs will be displayed.

TEXPAND
Displays the expanded version of the variance components table, the table of coefficients, the random effect predictions table, the table of fits and diagnostics, and the table of conditional means.

TSIMPLE
Displays the simple version of the variance components table, the table of coefficients, the random effect predictions table, the table of fits and diagnostics, and the table of conditional means.

TMETHOD
Displays the method table.

TFACTOR
Displays the name, number of levels, and the values for all categorical factors in your model.

TITERATION
Displays the iteration history table that shows the value of −2 log likelihood for each iteration of the variance components.

TVARIANCE
Displays the variance component estimates and the tests of whether the components are 0.

TCOVARIANCE
Display the asymptotic variance-covariance matrix of the variance component estimates.

TFIXEDEFFECT
Display tests of the fixed factors in the model.

TSUMMARY
Display S, R^2, and R^2(adj). The expanded version of the table adds AICc and BIC.

TCOEFFICIENTS
Displays the table of coefficients for fixed factors and covariates. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL
Displays the complete set of coefficients for categorical predictors.

NOFULL
Displays only the linearly independent coefficients.

TRANDOM
Displays the Best Unbiased Linear Predictions (BLUP) and tests of whether the BLUP are 0.

TEQUATION
Displays the marginal fitted equation. Minitab will display up to 50 equations. If you do not specify SINGLE or SEPARATE, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

SINGLE

If you want to view a single equation, rather than a separate equation for each factor level combination, use SINGLE.

SEPARATE

To view the separate equations, use SEPARATE.

TCEQUATION

Displays the conditional fitted equation. Minitab will display up to 50 equations. If you do not specify SINGLE or SEPARATE, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

SINGLE

If you want to view a single equation, rather than a separate equation for each factor level combination, use SINGLE.

SEPARATE

To view the separate equations, use SEPARATE.

TDIAGNOSTICS K

Displays a table of diagnostics for the marginal fits and residuals. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations.

TCONDITIONAL K

Displays a table of diagnostics for the conditional fits and residuals. K = 0 displays diagnostics for only unusual observations. K = 1 displays diagnostics for all observations.

TMEANS

Displays the table of least squares means.

Storage

RESIDUALS C

Stores the residuals for the marginal fits.

SRESIDUALS C

Stores the standardized residuals for the marginal fits.

TRESIDUALS C

Stores the deleted t residuals for the marginal fits.

FITS C

Stores the marginal fitted values.

CRESIDUALS C

Stores the conditional residuals.

CSRESIDUALS C

Stores the standardized residuals for the conditional fits.

CTRESIDUALS C

Stores the deleted t residuals for the conditional fits.

CFITS C

Stores the conditional fitted values.

HI C
Stores the leverages.

COOK C
Stores Cook's distance.

DFITS C
Stores DFITS.

COEFFICIENTS C
Stores the estimated coefficients.

BLUP C
Stores the best linear unbiased predictors.

XMATRIX *M*
Stores the design matrix for fixed effects terms.

ZMATRIX *M*
Stores the design matrix for random effects terms.

COVARIANCE *M*
Stores the variance-covariance matrix of the variance component estimates.

SMEANS C
Stores the least square means for the terms specified by MEANS.

VARCOMP C
Stores the estimates of the variance components in the same order as the terms appear after TERMS.

LOGLIKE C
Stores the −2 log likelihood values for all iterations.

FIXCOV *M*
Stores the variance-covariance matrix of the fixed parameters.

COMPARE: Session command for performing multiple comparisons of means

COMPARE C
Generates comparisons based on a model from a general linear model or from a mixed effects model for response C.

If you have a general linear model or a mixed effects model stored in the worksheet, you can use COMPARE to obtain multiple comparisons of means. Multiple comparisons of means allow you to examine which means are different and to estimate by how much they are different.

You have the following choices when using multiple comparisons:

- All pairwise comparisons or comparisons with a control
- Which means to compare

- The method of comparison
- How to display the results

Pairwise comparisons or comparison with a control

Use PAIRWISE when you do not have a control level but you would like to examine which pairs of means are different.

Use MCONTROL when you are comparing treatments to a control. When this method is suitable, it is inefficient to use the all-pairwise approach, because the all-pairwise confidence intervals will be wider and the hypothesis tests less powerful for a given family error rate.

Which means to compare

To specify which means to compare, enter terms from the model in the termslist for PAIRWISE or MCONTROL. If you have 2 factors named A and B, entering A B will result in multiple comparisons within each factor. Entering A * B will result in multiple comparisons for all level combinations of factors A and B.

The multiple comparisons method

You can choose from among four methods for both pairwise comparisons and comparisons with a control. Each method provides simultaneous or joint confidence intervals, meaning that the confidence level applies to the set of intervals computed by each method and not to each one individual interval. By protecting against false positives with multiple comparisons, the intervals are wider than if there were no protection.

Some characteristics of the multiple comparison methods are summarized below. "Conservative" in this context indicates that the true family error rate is less than the stated one.

Comparison method	Properties
Dunnett	Comparison to a control only, not proven to be conservative
Tukey	All pairwise differences only, not proven to be conservative
Bonferroni	Most conservative
Sidak	Conservative, but slightly less so than Bonferroni
Fisher's LSD	Does not control the family error rate

Display results

Minitab can present multiple comparison results in the following forms:

- TGROUPING displays the grouping information table that highlights the significant and non-significant comparisons.
- TMTEST displays the hypothesis test form which includes the adjusted p-values.
- GINTPLOT displays the interval plots of the confidence the intervals for the mean difference between two groups.

Conditions that restrict comparisons for general linear models

For general linear models, comparisons cannot be done in the following situations:

- Comparisons are not available for terms that contain or interact with random factors. Nesting is a form of interaction. For a model with random factors, you can do comparisons if you use a mixed effects model with the Restricted Maximum Likelihood estimation method (REML). Suppose the model contains A B C A*B, where both A and C are random, and B is fixed. Because A*B is a term in the model, no multiple comparison results for B are available even though B is a fixed factor. However, if the model does not include A*B, multiple comparison results are available for B.

  There is a special case for a balanced design that has two factors. Suppose the model contains A B A*B, where A is random, and B is fixed. Even though A*B is in the model, you can perform multiple comparisons for B.

- Comparisons is disabled if you choose (**1, 0**) coding and the model is non-hierarchical. To enable **Comparisons** for this case, choose (**−1, 0, +1**) coding or specify a hierarchical model.

Commands

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PAIRWISE ***[termlist]***

Specifies pairwise comparisons using the terms in the termlist. The terms in the termlist must be in the model.

You must use at least one of the following sub-subcommands with PAIRWISE:

TUKEY

BONFERRONI

SIDAK

FISHER

MCONTROL ***[termlist]***

Specifies comparisons with a control using the terms in the termlist. The terms in the termlist must be in the model.

You must use at least one of the following sub-subcommands with PAIRWISE:

DUNNETT

BONFERRONI

SIDAK

FISHER

LEVELS C K...C K

You must use LEVELS with MCONTROL. Use LEVELS to provide a factor (C) with a control level (K). You must specify a control level for each factor that you list in MCONTROL. If these levels are text or date/time, enclose each with double quotation marks.

ALTERNATIVE K

ALTERNATIVE specifies the alternative hypothesis for comparisons to the control level.

Value of K	Alternative hypothesis
−1	Determines whether the treatment means are less than the control mean
0	Determines whether the treatment means are equal to the control mean

Value of K	Alternative hypothesis
1	Determines whether the treatment means are greater than the control mean

GINTPLOT

Displays interval plots that represent the confidence interval for the mean difference between two groups. A separate graph is created for each comparison.

TGROUPING

Displays the grouping information table for each comparison. This table highlights the significant and non-significant comparisons for each selected comparison method.

TMTEST

Displays the multiple comparison test table. This table displays the hypothesis test form of the comparison output, which includes the differences of the means, the numeric values for the confidences intervals, and the adjusted p-values.

MANOVA: Session command for performing a general MANOVA

MANOVA C...C = C...C

Performs a general MANOVA. The columns specified before the equal sign are the responses, and the columns after the equal sign are the factors.

Use general MANOVA to perform multivariate analysis of variance (MANOVA) with balanced and unbalanced designs, or if you have covariates. This procedure takes advantage of the data covariance structure to simultaneously test the equality of means from different responses.

Calculations are done using a regression approach. A full rank design matrix is formed from the factors and covariates and each response variable is regressed on the columns of the design matrix.

Factors can be crossed or nested, but they cannot be declared as random. You can work around this restriction by specifying the error term to test model terms. Covariates can be crossed with each other or with factors, or nested within factors. You can analyze up to 50 response variables with up to 31 factors and 50 covariates at one time.

With the MANOVA subcommand, you can specify model terms for a custom multivariate test and designate the error term in Error. Minitab performs four multivariate tests (see the MANOVA subcommand) for those terms. This option is most useful when you have factors that you consider as random factors. Model terms that are random or that are interactions with random terms may need a different error term than general MANOVA supplies. You can determine the appropriate error term by entering one response variable with General Linear Model, choose to display the expected mean square, and determine which error term was used for each model terms.

If you specify an error term, it must be a single term that is in the model. This error term is used for all requested tests. If you have different error terms for certain model terms, enter each separately and exercise the general MANOVA dialog for each one. If you do not specify an error term, Minitab uses MSE.

For more information, go to How to specify the model for GLM on page 1250 and How to enter data for ANOVA and GLM on page 1241.

Model

MANOVA ***[termlist [/ errorterm]]***

Performs four multivariate tests—Wilks' test, Lawley-Hotelling test, Pillai's test, and Roy's largest root test—for each term in the model. If you include a termlist, MANOVA does the four multivariate tests for each term listed.

If you specify an errorterm, it must be a single term that is in the model. MANOVA then uses this errorterm in all tests. If you do not specify an errorterm, Minitab uses the error associated with MSE, as in the univariate case.

All four tests are based on two SSCP (sums of squares and cross products) matrices: H = the hypothesis matrix and E = the error matrix.

If an error term is not specified on MANOVA, then the adjusted SSCP matrix is used for H and the SSCP matrix associated with MSE is used for E. If an error term is specified on MANOVA, the sequential SSCP matrices associated with H and E are used. Using sequential SSCP matrices guarantees that H and E are statistically independent.

COVARIATES C...C

The columns listed are used as covariates. If COVARIATES is used, it must be the first subcommand. This restriction is needed to allow proper error checking. You may have up to 50 covariates.

Options

WEIGHTS C

Performs a weighted least squares fit. The weights must be greater than or equal to zero. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are calculated by $(\mathbf{X'WX})^{-1}(\mathbf{X'WY})$.

This is equivalent to minimizing the weighted SS Error:

$$\Sigma w_i(Y - \hat{Y})^2$$

PREDICT

Computes the fitted Y's, or $\hat{Y}$ for given values of the predictors. PREDICT displays a table that contains the fitted Y's, standard errors of the fitted Y's, a 95% confidence interval, and a 95% prediction interval. E...E may be a list of values, one for each predictor, a list of columns, one for each predictor, or a mix of values and columns.

The prediction interval computed by PREDICT assumes a weight of 1. If you used the WEIGHT subcommand with values other than 1, you should adjust the prediction interval values manually.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PFITS C

Stores the fits.

PSDFITS C

Stores the estimated standard errors of the fits.

CLIMITS C C, ..., C C

Stores the lower and upper confidence limits.

PLIMITS C C, ..., C C

Stores the lower and upper prediction limits.

TEST *termlist / errorterm*

Use TEST to specify your own tests. Termlist is a list of terms in the model. List them exactly as on the MANOVA subcommand. Error term is the term to be used as the denominator for the F-test. This can be a term in the model, a linear combination of terms in the model, or MSE (denoted by the word ERROR). Only available via command line.

TOLERANCE K [K]

Use TOLERANCE to force Minitab to keep a predictor in the model which is either highly correlated with another predictor or which is nearly constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

K	Description
First K	1 – R-squared, where R-squared is the value resulting from regressing one predictor on the remaining predictors. The default is 1E – 18. If (1 – R-squared) is $\leq$ (1E – 18), the predictor is removed from the equation.
Second K	The parameter for forcing in a variable which is nearly constant. The default is 2E – 21. If the coefficient of variation of a predictor (the standard deviation / the mean) is $\leq$ square root (2E – 21), then the predictor is considered essentially constant and is removed from the equation.

Both Ks must be positive numbers.

Graphs

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n). Optionally, specify a column C that defines the order.

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C

Displays a separate graph for the residuals versus each specified column.

Results

MEANS *termlist*

Displays a table of adjusted means (sometimes called least squares means) corresponding to each term listed. Terms listed on MEANS must also be listed on MANOVA. If more than one MEANS subcommand is specified, only the last one is used.

EMS

Displays a table that contains expected mean squares, estimated variance components, and the error term (the denominator) used in each exact F-test. If there is no exact F-test for a term, the expected mean squares allow you to determine how to construct an approximate F-test using the subcommand TEST.

The estimates of the variance components are the usual unbiased analysis of variance estimates. They are obtained by setting each calculated MS equal to its EMS. This gives a system of linear equations in the unknown variance components. This system is then solved. Unfortunately, this method can result in negative estimates, which should be set to zero. Minitab, however, prints the negative estimates because they sometimes indicate that the model being fit is inappropriate for the data.

Terms that are fixed do not have variance components estimated.

BRIEF K

Controls the amount of output with a value in K. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage and displays the following error messages, warnings, prompts, and notes; graphs; WRITE to the screen.
1	Minitab displays the ANOVA table.
2 (default)	K = 1, and Minitab displays the table of factor levels, coefficients for terms involving covariates, and unusual observations.
3	K = 2 output, and Minitab displays all the coefficients.

SSCP

Displays the hypothesis matrix, H, corresponding to each term specified by MANOVA, and the error matrix, E.

EIGEN

Displays a table containing the eigenvalues and eigenvectors for the (nonsymmetric) matrix, E**-1 H. These are the eigenvalues that are used to calculate the four MANOVA tests. A separate table is printed for each term specified on MANOVA.

Note If an eigenvalue is repeated, then the corresponding eigenvectors are not unique. In this case, the eigenvectors Minitab displays and those in books or other software might not agree. However, the MANOVA tests are always unique.

PARTIAL

Displays a matrix of partial correlations. These are the correlations among the residuals or, equivalently, the correlations among the responses conditioned on the model. The formula for the matrix is W**-.5 E W**-.5, where E is the error matrix and W has the diagonal of E as its diagonal and 0's elsewhere.

NOUNIVARIATE

Suppresses the univariate output. Only the multivariate output is displayed.

INTERACT *terms*
Used with SMEANS to store means for two-way interactions. List the model terms for which you calculate means for level pairs. The means will be calculated by the order of terms and levels.

SMEANS C...C
Stores means for two-way interactions, main effects, and the overall mean. Use with INTERACT to store means for two-way interactions. Specify a storage column for each response variable.

Storage

FITS C...C
Stores fitted values, using one column for each response.

RESIDUALS C...C
Stores residuals, using one column for each response variable. Residual = (response – fit).

SRESIDUALS C...C
Stores the standardized residuals, using one column for each response variable.

TRESIDUALS C...C
Stores the deleted Studentized residuals, using one column for each response variable.

HI C
Stores leverages.

COOKD C...C
Stores Cook's distance.

DFITS C...C
Stores the DFITS (also called DFFITS), using one column for each response.

XMATRIX *M*
Stores the design matrix corresponding to your model in M.

COEFFICIENTS C...C
Stores the coefficients for a model, using one column for each response. These are the same coefficients that are printed under BRIEF 3. They correspond to the design matrix stored by XMATRIX. Thus, if M1 contains the design matrix and C1 the coefficients, then M1 times C1 gives the fitted values.

NESTED: Session command for performing a fully-nested ANOVA

NESTED C...C = C...C
Analyzes fully nested (hierarchical) designs. All factors are assumed to be random. Designs can be unbalanced. In this case, F and p-values are not displayed. You can calculate these yourself using the expected mean squares that are displayed.

You can analyze up to 50 response variables and up to 9 factors on one NESTED command. The columns specified before the equal sign are the responses, and the columns after the equal sign are the factors. You can also have replicates. One row then corresponds to one observation, giving the value of each response and the level of each

factor for that observation. Factor levels can be any real numbers. They do not need to be consecutive or in any special order.

VARTEST: Session command for performing an equal variances test

VARTEST C...C

If you do not specify UNSTACKED, VARTEST performs tests for equal variances with all response data in one C and factor level information in additional columns.

If you do specify UNSTACKED, VARTEST performs tests for equal variances with data from each factor level in a different column C...C.

UNSTACKED

Indicates that data for each factor level are in different columns.

Options

CONFIDENCE K

Specifies the confidence level for the Bonferroni simultaneous confidence intervals and also the significance level (denoted by α or alpha) for the multiple comparison intervals, and the tests. The default is 95, which corresponds to a confidence level of 95% and an $\alpha = 1 - (95 / 100) = 0.05$.

USEBARTLETT

Specifies to use the test based on the normal distribution instead of the multiple comparisons method and Levene's method. If you have only 2 factor levels, then Minitab performs the F-test. If you have 3 or more factor levels, then Minitab performs Bartlett's test.

The F-test and Bartlett's test are accurate only for normally distributed data. Any departure from normality can cause these tests to yield inaccurate results. However, if the data conform to the normal distribution, then the F-test and Bartlett's test are typically more powerful than either the multiple comparisons method or Levene's method.

Graphs

GINTERVAL

Displays a graphical summary that includes the multiple comparison intervals for standard deviations and test results by default. The multiple comparison intervals are not confidence intervals of the individual standard deviations. The multiple comparison intervals are only useful for comparing the population standard deviations. If USEBARTLETT is used, then GINTERVAL displays Bonferroni confidence intervals of the standard deviations.

GINDPLOT

Displays individual value plot.

GBOXPLOT

Displays a boxplot.

Results

NODEFAULT
Minitab displays all output tables by default. You do not have to enter the subcommands. If you enter NODEFAULT, then each table is only displayed if you enter the specific subcommand.

TMETHOD
Displays the method table which includes the null hypothesis, the alternative hypothesis, and the significance level (denoted by α or alpha).

TBONFERRONI
Displays the table of Bonferroni simultaneous confidence intervals for the standard deviations of each factor level.

TTEST
Displays the test table which includes the p-values for the hypothesis tests.

Storage

STDEVS C
Stores the standard deviation of each factor level in column C.

VARIANCES C
Stores the variance of each factor level in column C.

SBONFERRONI C C
Stores Bonferroni confidence limits for the standard deviation of each factor level. The lower limits are stored in the first column C and the upper limits is stored in the second column C.

SMCI C C
Stores the limits of the multiple comparisons intervals for each factor level. The lower limits are stored in the first column C and the upper limits is stored in the second column C.

SFPVALUE C
Stores the p-value for Bartlett's test (or the F-test if there are only 2 factor levels).

SLPVALUE C
Stores the p-value for Levene's test.

SMPVALUE C
Stores the p-value for the multiple comparison test.

INTPLOT: Session command for creating an interval plot

INTPLOT C...C

INTPLOT (C...C) * C
Use to plot means and confidence intervals for one or more variables. An interval plot illustrates both a measure of central tendency and the variability of the data.

By default, Minitab displays confidence intervals, but you can change the display type to standard error bars using the INTBAR subcommand.

The data must be numeric or date/time. The categorical grouping data can be numeric, date/time, or text.

INTPLOT C...C displays a separate interval plot for each graph variable.

INTPLOT (C...C) * C displays a separate graph for each C on the left, with an interval for each category of the C on the right.

The following table shows a few of the possible ways to generate interval plots.

Command language	What the command language displays
`INTPLOT 'Sales' 'Advertis'.`	Two graphs: interval plots of Sales and of Advertis
`INTPLOT 'Sales'*'Year'.`	A graph with intervals representing yearly sales

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot on page 1106

INTLAB: Session subcommand for labeling interval bar endpoints on an interval plot on page 1112

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot on page 1124

MEDLAB: Session subcommand for labeling medians on a boxplot on page 1128

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

BAR: Session subcommand for representing data values with bars on page 1076

CMEAN: Session subcommand for connecting means with lines on a boxplot on page 1083

CMEDIAN: Session subcommand for connecting means with lines on a boxplot on page 1084

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot on page 1109

INTBAR: Session subcommand for displaying a vertical line with horizontal lines at the endpoints of the confidence interval for the mean on page 1110

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot on page 1126

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot on page 1127

SOFFSET: Session subcommand for offsetting points from the center on page 1155

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

MAIN: Session command for creating a main effects plot

MAIN C...C

Draws a main effects plot where the points in the plot are the raw data means at the various levels of each factor. A reference line is drawn through the plot representing the grand mean of the response data. You can use factors from designs generated by commands such as GLM on page 228 and ANOVA on page 225.

MAIN generates a plot that uses data means. After you have fit a model, use FACPLOT on page 277 to generate plots that use fitted means.

Data means are the raw response variable means for each factor level combination whereas fitted means use least squares to predict the mean response values of a balanced design. Therefore, the two types of means are identical for balanced designs but can be different for unbalanced designs. While you can use raw data with unbalanced designs to obtain a general idea of which main effects may be evident, it is generally good practice to use the fitted means to obtain more precise results.

Enter the factors on the main command line. The factor columns may be numeric or text, and may contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on.

RESPONSE C

Use RESPONSE to enter raw response data. Then, the main effects are calculated based on the means of the data.

YMIN K

Specifies the y-axis minimum value. Making the minimum and maximum values the same is very useful when you are comparing several main effects plots of related data.

YMAX K

Specifies the y-axis maximum value. Making the minimum and maximum values the same is very useful when you are comparing several main effects plots of related data.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

INTERACT: Session command for creating an interactions plot

INTERACT C...C
Draws a single interaction plot if 2 factors are entered on the command line, or a matrix of interaction plots if more than 2 factors are entered. You can use factors from designs generated by commands such as GLM on page 228 and ANOVA on page 225.

INTERACT generates a plot that uses data means. After you have fit a model, use FACPLOT on page 277 to generate plots that use fitted means.

Data means are the raw response variable means for each factor level combination whereas fitted means use least squares to predict the mean response values of a balanced design. Therefore, the two types of means are identical for balanced designs but can be different for unbalanced designs. While you can use raw data with unbalanced designs to obtain a general idea of which main effects may be evident, it is generally good practice to use the fitted means to obtain more precise results.

Enter the columns containing the factors as the arguments to the main command. The factor columns may be numeric or text, and may contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on.

You must use the RESPONSE subcommand with a column argument. When you specify RESPONSE and a response column, Minitab goes through each two factor pair in sequential order and calculates the means of all the responses for each unique combination of values in the two factors.

RESPONSE C

Specifies the column containing the response data.

The RESPONSE version of the interaction plot does not require any model to be fit. It is useful for preliminary exploratory analysis in order to get a preliminary look at which two factor interactions appear to be significant. It can also be used after a model has been fit, provided the design is balanced.

FULL

Specifies to use the full interaction matrix when more than two factors are specified instead of the default display of the upper right portion of the matrix. In the full matrix, the transpose of each plot in the upper right displays in the lower left portion of the matrix. The full matrix takes longer to display than the half matrix.

YMIN K

Specifies the y-axis minimum value. Making the minimum and maximum values the same is very useful when you are comparing several main effects plots of related data.

YMAX K

Specifies the y-axis maximum value. Making the minimum and maximum values the same is very useful when you are comparing several main effects plots of related data.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

DOE

Screening Designs

DSDESIGN: Session command for creating a definitive screening design

DSDESIGN K

Creates a definitive screening design with K factors. You can then use REPLICATES to replicate points and BLOCKS to add block.

The design summary and design table can be displayed. The design table is displayed under BRIEF 3. Factors are denoted by the letters A, B, C, D, E, F, G, H, J, K, ... , vv, ww, xx.

Note The letter I is not used.

Quotation marks around column names are required. If you use a subcommand more than one time, DSDESIGN uses the last instance of the subcommand.

REPLICATES K

Replicates each point K times.

The following example generates a definitive screening design with 4 factors and 2 replicates. Thus, there are a total of 26 runs in this experiment.

```
DSDESIGN 4;
  REPLICATES 2.
```

BLOCKS K

Specifies the number of blocks in K. If your design includes blocks, and the number of blocks equals the number of replicates, each replicate is a separate block. You can include multiple replicates in a block. Each block has the same number of replicates.

LEVELS K K, ..., K K

With LEVELS K K, ..., K K, you list one pair of arguments for each factor. The first K should be the low level, and the second K should be the high level. Minitab calculates the center points as the value that is half way between the low and high levels.

The following example stores 0 for the low level of A, 10 for the high level, and 5 for the center point; 100 for the low level of B, 102 for the high level, and 101 for the center point; .2 for the low level of C, .3 for the high level, and 2.5 for the center point.

```
DSDESIGN 3;
 LEVELS  0 10 100 102 .2 .3.
```

With LEVELS K K, ..., K K, you can set levels to be text data. The arguments that define text data must be enclosed in double quotation marks. For example, if factor B is whether a catalyst is present or absent, you can use the following:

```
DSDESIGN 3;
  LEVELS  0 10 "absent" "present" .2 .3.
```

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block, and then the blocks are randomized.

K is the base for the random number generator and can take any positive integer as an argument. To always get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

FNAME K...K

Specifies the names of the factors. You must enter as many arguments as number of factors. The arguments must be enclosed in double quotation marks.

BRIEF K

Controls the amount of output.

You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1 or 2	Minitab displays a summary of the design.
3	In addition to the K = 1 or K = 2 output, Minitab displays the design table.

PBDESIGN: Session command for creating a Plackett-Burman design

PBDESIGN K [K]

Creates Plackett-Burman designs. Plackett-Burman designs are a class of resolution III, 2-level fractional factorial designs that are often used to study main effects. In a resolution III design, main effects are aliased with 2-way interactions.

The two constants on PBDESIGN specify a design. The first K is the number of factors. The second K is the number of runs. Factor columns can be text or numeric.

To see the designs that Minitab generates, go to Plackett-Burman designs on page 1259. A design is based on the number of runs, from 8 to 48 and always a multiple of 4. Therefore, the second K on PBDESIGN must be a multiple of 4. The number of factors must be less than the number of runs. If you don't include the second K, Minitab sets the number of runs to the smallest possible value for the specified number of factors.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTPYE, and XMATRIX.)

CENTER K

Adds up to 50 center points to the design. If the design is blocked, the center points are divided equally between the blocks. If K is not a multiple of the number of blocks, then each of the last few blocks has one less point than the other blocks.

REPS K

Specifies up to 50 replicates of each point. REPS does not replicate center points.

The following example generates a Plackett-Burman design that has 2 factors and 8 base runs. Thus, there are a total of 16 runs in this experiment.

```
PBDESIGN 2 8;
 REPS 2.
```

RBLOCK

Specifies blocks on replicates. Each set of replicate points are placed in a separate block. No arguments are taken.

LEVELS

Use LEVELS to specify values for the levels of each factor. The level values can be text or numeric. The default values are -1 for the low level, +1 for the high level, and 0 for a center point.

The two ways to use LEVELS are as follows.

LEVELS K K, ..., K K

With LEVELS K K, ..., K K, you list one pair of arguments for each factor. The first K should be the low level and the second K should be the high level. When you add center points with CPBLOCK, Minitab calculates the center points as the value that is half way between the low and high levels. The following example stores 0 for the low level of A, 10 for the high level, and 5 for the center point. 100 for the low level of B, 102 for the high level, and 101 for the center point; .2 for the low level of C, .3 for the high level, and 2.5 for the center point.

```
PBDESIGN 2 8;
  CENTER 2;
 LEVELS  0 10 100 102;
 DESIGN.
```

With LEVELS K K, ..., K K, you can set levels to be text data. The arguments that define text data must be enclosed in double quotation marks. In the following example, factor B is whether a catalyst is present or absent. Because you use CENTER 2 to add center points with a text factor, Minitab creates pseudo center points.

```
PBDESIGN 2 8;
  CENTER 2;
  RBLOCK;
  LEVELS  0 10 "absent" "present";
  DESIGN.
```

LEVELS C

With LEVELS C, the level values must be numeric. Put the level values in a column, in the same order as if they were listed on the subcommand. Then, specify the level column on the LEVELS subcommand.

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output for the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

SCREEN: Session command for analyzing a screening design

SCREEN

Analyzes a screening design with a continuous response. Fits definitive screening and Plackett Burman models, which are produced by DSDESIGN on page 258 and PBDESIGN on page 259. Each factor must have two levels. Also, you can have covariates and a blocking factor. You can specify a single response. Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, Minitab uses the first entry (row 1) for the low level of the factor and the other entry for the high level. To change this default order, use the LEVEL subcommand. When you include center points in the model, Minitab tests for curvature.

You can display effects plots and residual plots. You can store the residuals, fitted values, and many other diagnostics for further analysis. You can use extra text on the command and subcommand lines only after a #. Quotation marks around column names that are one word long are optional.

For information about how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

RESPONSE C

Specifies the column that contains the response variable in C. The column must be numeric or date/time.

BLOCK C (optional)

Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

FACTOR C...C (required)

Specifies the factors in columns C...C. The columns must have the same number of rows that the response column has.

COVARIATES C...C

Specifies columns that include covariates in C...C. Include the covariate columns with terms to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

LEVELS K K,..., K K

LEVELS C

Specifies the uncoded factor levels for each factor, with K for the low level and K for the high level, or in C. For each factor, the list or column must contain two values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all of the levels are continuous or all of the levels are categorical.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

TERMS *termlist*

Specifies the terms to include in the model. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

Options

WEIGHTS C

Specifies a numeric column (C) that contains weights to use in weighted regression analysis. The same weights are used for all responses. If different responses require different weights, then SCREEN will need to be run as many times as there are different weights.

Usually, the weights are the reciprocal of the variances of the responses. The leverages stored by HI are affected by weights.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

BOXCOX [K]

Transforms each variable or group to follow a normal distribution using the lambda value in K. K must be between –5 and +5. If you do not specify K, then Minitab finds the optimal lambda. By default, Minitab rounds the optimal value.

To use the exact lambda, change the setting in **File** > **Options** > **Linear Models** > **Display of Results**.

Minitab cannot calculate the optimal lambda for stepwise regression. So you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

NOROUNDING

Specifies no rounding of calculated lambda parameter.

MEANS *termlist*

Calculates the least squares means for the terms that you specify in the termlist. You can use the following commands to display and store the least squares means.

TMEANS

Displays the least squares means for the specified terms.

SMEANS

Stores the least squares means for the specified terms.

BSMEANS

If you transformed the data, stores the least squares means for the specified terms.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors.

Requires that Minitab retain a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

Stepwise regression

`FINFORMATION`, `STEPWISE`, `FORWARD`, and `BACKWARD` perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value

of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED
Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION
Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE
Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD
Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD
Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K
Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K
Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*
Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*
Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL
Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL
Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY
Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. K can only be 1 or 2. When K = 2, Minitab includes square terms as well as 2-factor interactions. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]

Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C

Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

The following subcommands specify what is displayed in the results.

NODEFAULT

Specifies that no default tables or graphs are displayed.

TEXPAND

Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE

Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TMETHOD

Displays the method table.

TMSDETAILS

Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL

Displays the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

NOFULL

Does not display the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

TEQUATION

Displays the regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TCOEFFICIENT

Displays the table of coefficients. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

Displays the summary of model table, which includes the R^2 statistics and S. The expanded version of the table includes the prediction sum of squares (PRESS), the corrected Akaike's Information Criterion (AICc) and the Bayesian Information Criterion (BIC).

TANOVA

Displays the ANOVA table. The expanded version of the table includes the sequential sums of squares and the percent contribution to the total sum of squares.

TDIAGNOSTICS [K]

Displays the table of diagnostics. The expanded version of the table includes the standard error of the fit, the confidence interval for the fit, the deleted residual, the leverage, Cook's D, and DFITS.

Value of K	Diagnostics that are displayed
0	Only for unusual observations
1	For all observations

TALIAS [K]

Displays the alias table up to order K. One line is displayed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, then Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to three-way interaction for 7 to 10 factors, and two-way interactions for more than 10.

TMEANS

Displays the table of least squares means for all combinations of the factor levels that you specify in MEANS.

Storage

The following subcommands specify what is stored in the worksheet.

FITS C
Stores the fitted values.

RESIDUALS C
Stores the residuals (response – fit) in C.

SRESIDUALS C
Stores the standardized residuals.

TRESIDUALS C
Stores the deleted Studentized residuals in C.

HI C
Stores the leverage values.

COOK C
Stores Cook's distance values, using one column for each response.

DFITS C
Stores DFITS (also called DFFITS), using one column for each response.

COEFFICIENTS C
Stores the coefficients that are included in the coefficients table in C.

The coefficients are based on coded levels (–1, 0, +1), not on the actual levels in the data set. A coefficient is always half the value of the corresponding effect.

XMATRIX *M*
Stores the design matrix. The first column is a column of ones for the constant term. Next, there is one column for each covariate. Next, if the design was blocked into k blocks, there are (k – 1) columns. Next, there is one column for each factor, using –1 for the low level, 0 for a center point, and +1 for the high level. Next, there is one column for each interaction in the model. The column for an interaction is the product of all the columns that correspond to the factors that are contained in that interaction.

If FFACTORIAL or VFACTORIAL removed some of the interactions that you specified in your model, the interactions are not included in the stored design matrix. The columns in XMATRIX match the coefficients that are displayed and stored.

BCRESP C
Stores the Box-Cox transformation of the response.

BFITS C
Stores the back-transformed fits after the Box-Cox transformation in C.

BSMEANS C
If you use a Box-Cox transformation, BSMEANS stores the means for the back-transformed response in C.

SMEANS C
Stores the least squares means from MEANS in C.

If you use a Box-Cox transformation, SMEANS stores the means for the transformed response in C.

BSCREEN: Session command for analyzing a screening design with a binary response

Required commands

To perform the analysis, use the main command and at least these accompanying subcommands.

BSCREEN

Analyzes a screening design with a binary response.

Fits definitive screening and Plackett Burman models, which are produced by DSDESIGN on page 258 and PBDESIGN on page 259. Each factor must have two levels. Also, you can have covariates and a blocking factor.

Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

This analysis provides diagnostic plots, goodness-of-fit tests, and other diagnostic measures so you can assess the validity of your model. For more information, go to Generalized linear model diagnostics and residual analysis on page 1260.

Minitab excludes all observations with missing values on either the response variable or any of the predictors from all calculations.

RESPONSE C C

Enters the response data as a column that contains the number of events first and then a column that contains the number of trials. The columns must have the same number of rows. For examples, go to Entering data for response variables on page 1225.

FACTOR C...C

Specifies the factors in columns C...C. The columns must have the same number of rows as the response columns.

TERMS *termlist*

Specifies the terms to include in the model. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

LEVELS K K,..., K K or LEVELS C

Specifies the uncoded factor levels for each factor. In a pair of levels, the first K is the low level and the second K is the high level. In a pair of column rows, the first row is the low level and the second row is the high level. For each factor, the list or column must contain two values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all of the levels are continuous or all of the levels are categorical.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

Optional subcommands

Use these subcommands and the subcommands in the following sections to change the default analysis.

EVNAME K
Specify the name of the event with K. Enclose text values in double quotation marks.

LOGIT
Use the logit link function.

NORMIT
Use the normit link function.

GOMPIT
Use the gompit link function. The gompit link function is also called the complementary log-log link function.

BLOCKS C
Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

COVARIATES C...C
Specifies columns that include covariates in C...C. Include the covariate columns with TERMS to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

Options subcommands

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K
Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SEQDEVIANCE
Use sequential deviances for the tests of significance in the ANOVA table. The order of the terms in TERMS affects the statistical significance. The default is to use adjusted deviances.

PEARSON
Specifies the use of Pearson residuals. The default is to use deviance residuals.

WEIGHT C
Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} (\mathbf{X}' \mathbf{W} \mathbf{Y})$

HGROUP K

Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test with K. The default is 10.

MEANS *termlist*

Calculates the fitted mean probability for the terms that you specify in the termlist. You can use the following commands to display and store the probabilities.

TMEANS

Displays the fitted mean probability for the specified terms.

SMEANS C

Stores the fitted mean probability for the specified terms.

ITERATION K

Specifies the maximum number of iterations for the optimization algorithm that estimates the coefficients. The default is 50.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors. By default, K = 1E-12.

The tolerance value determines when Minitab retains a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

CTOLERANCE K

Specifies the convergence criterion for the algorithm that estimates the coefficients. By default, K = 1E-8.

START C

Specifies a column that contains starting values for the coefficients in the optimization algorithm. The column begins with the estimate of the constant, then follows the order of the model after TERMS.

Stepwise regression subcommands

`FINFORMATION`, `STEPWISE`, `FORWARD`, and `BACKWARD` perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example,

a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY
Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS
Specifies that both categorical and continuous terms have to be hierarchical.

ATEND
Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS
Specifies that the model is hierarchical at every step.

SINGLE
Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE
Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs subcommands

RTYPE K
Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GEFFECTS
Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF
Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Displays residuals versus fitted values.

GORDER [C]

Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]

Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C

Displays the residuals versus the variables in C. Typically, you enter the columns for the predictors.

Results subcommands

NODEFAULT

Hides default tables and graphs that you do not include the subcommand to display.

TEXPAND

Displays the expanded versions of the tables that have them. The expanded Analysis of Variance table adds the sequential deviances and percent contributions for each term. The Coefficients table adds confidence intervals for the coefficients. The Fits and Diagnostics table adds the standard error of the fit, the 95% confidence interval for the fit, the deleted residual, the HI, the Cook's D, and the DFITS.

TSIMPLE

Displays the simple versions of the Analysis of Variance table, the Coefficients table, and the Fits and Diagnostics table.

TMETHOD

Displays a table of information about the method and options.

TMSDETAILS

Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model.

If you do not specify FULL or NOFULL, then Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL

Specifies to display the coefficients, p-values, and model summary statistics for each step of the procedure.

NOFULL

Hides these statistics.

TRINFO

Displays the number of events, non-events, and the total. The table also indicates which response value is the reference event.

TSTEP

Displays the deviance at each iteration of the coefficients.

TCOEFFICIENTS

Displays the coefficients and the p-values for the Wald normal approximation tests.

If you do not specify FULL or NOFULL, Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Specifies to show the coefficients for the reference level when you use (0, 1) coding for categorical predictors.

NOFULL

Hides these statistics.

TODDS

Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model.

INCREMENT K...K

Specifies the increment for the odds ratio for the predictor with K. For example, if the predictor is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram. Enter a value for each numeric factor in the same order that the factors follow the FACTOR subcommand. Then, enter a value for each covariate in the order that the covariates follow the COVARIATE subcommand. Enter 1 for a factor or covariate to use the same units as the data.

TSUMMARY

Displays the statistics that evaluate model fit, including deviance R^2.

TGOODNESS

Displays the Pearson and deviance goodness-of-fit tests. For the binomial distribution models, also shows the Hosmer-Lemeshow goodness-of-fit test.

TDEVIANCE

Displays the ANOVA table.

TEQUATION

Displays the regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TMEANS
Displays the table of fitted mean probabilities for all combinations of the factor levels that you specify in MEANS.

THOSMER
Displays the table of observed and expected values for the Hosmer-Lemeshow test.

TASSOCIATION
Displays the number of concordant and discordant pairs, the Somers' D, the Goodman-Kruskal Gamma, and Kendall's Tau-a.

TDIAGNOSTICS [K]
Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

TALIAS [K]
Displays the alias table up to order K. One line is displayed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, then Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to three-way interaction for 7 to 10 factors, and two-way interactions for more than 10.

Storage subcommands

FITS C
Stores the event probabilities in C.

RESIDUALS C
Stores the deviance residuals by default in C. Stores the Pearson residuals if you use the subcommand PEARSON.

SRESIDUALS C
Stores the standardized deviance residuals by default in C. Stores the standardized Pearson residuals if you use the subcommand PEARSON.

TRESIDUALS C
Stores the deleted deviance residuals by default in C. Stores the deleted Pearson residuals if you use the subcommand PEARSON.

COOK C
Stores the Cook's distances in C.

DFITS C
Stores the DFITS in C.

HI C
Stores the leverages in C.

SADJUSTEDRESP C
Stores the final adjusted response values in C.

SIWEIGHTS C
Stores the internal weights used to estimate the model parameters in C.

SECOEFFICIENTS C
Stores the standard errors of the estimated coefficients in C.

COEFFICIENTS C
Stores the estimated coefficients in C.

EFFECTS C
Stores the effects that are included in the output in C.

Note Effects are not displayed or stored for the constant, covariates, or blocks.

SMEANS C
Stores the fitted mean probability for the specified terms in C.

DBETA C
Stores the change in the estimated regression coefficients (delta b) when you delete all observations with a particular factor/covariate pattern in C. Use the delta b values to detect observations with a strong influence on the coefficients.

DSBETA C
Stores the change in the estimated regression coefficients when you delete all observations with a particular factor/covariate pattern based on the standardized Pearson residual in C.

DCHISQUARE C
Stores the change in the chi-square statistic when you delete all observations with a particular factor/covariate pattern in C. Observations that are poorly fit by the model have high delta chi-square values.

DDEVIANCE C
Stores the change in the deviance when you delete all observations with a particular factor/covariate pattern in C.

XMATRIX *M*
Stores the design matrix in M.

XPWXINVERSE *M*
Stores the X'WX inverse matrix of the estimated coefficients in M.

Plot Factorial Design

FACPLOT: Session command for creating a factorial plot

FACPLOT C

FACPLOT "T|K"
Displays factorial plots for the model associated with the response variable in the specified column. To create this graph, you need to specify a model. To produce an interaction plot, you must have two or more variables.

Displays factorial plots for the model that you specify. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following code: `FACPLOT 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `FACPLOT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

To produce an interaction plot, you must have two or more variables. These plots can illustrate how a response variable relates to one or more variables. You can produce two types of factorial plots.

- Use a main effects plot to compare fitted means or data means across levels of one or more variables.
- Use an interactions plot to assess the effects of one variable at levels of another variable. An interactions plot displays the fitted means for a level of a variable with the level of a second variable held constant.

Although you can use these plots to display the effects, be sure to evaluate significance by looking at the effects in the analysis of variance table.

FACTORS C...C

Specifies the variables to display in both the main effect and interaction plots. You must specify 2 or more variables to generate an interaction plot.

GMAIN

GMAIN displays the main effect plot.

GINT

GINT displays the interactions plot.

FULL

Use FULL to display both interaction plots for each pair of categorical variables, such as Predictor 1 by Predictor 2 and Predictor 2 by Predictor 1. Sometimes you can identify interactions more easily when you examine both plots.

LOWER

Use LOWER to display one version of each interaction plot.

GMODEL (default)

GMODEL display plots for only terms that are in the model and listed in FACTORS. For purposes of a main effect plot, a variable is in the model if the model includes any terms that involve only that variable, e.g. A, A*A, or A*A*A. For purposes of an interaction plot, the interaction of two variables is in the model if the model includes any terms that involve both variables, e.g. A*B, A*B*B.

GALL

GALL displays plots for all terms involving variables on the FACTORS subcommand.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

Factorial Designs

FDESIGN: Session command for creating a general full factorial design

FDESIGN K...K

FDESIGN creates a general full factorial design. The values of K specify the number of levels of each factor. The values of K must be integers that are greater than or equal to 2. The model generated is always full. To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.)

FDESIGN is similar to FFDESIGN, but FDESIGN has arbitrary levels and no fractional designs.

REPS K

Replicates each point K times. Specifies up to 50 replicates of each point. REPS does not replicate center points.

The following example generates a full factorial that has 3 factors and 2 observations per cell. Thus, there are a total of 48 runs in this experiment.

```
FDESIGN 4 3 2;
  REPS 2.
```

BLOCK C

Specifies a blocking variable column (C). You can specify up to 64 blocks. The blocking variable column can contain numeric data or text data, and can contain any value. The levels do not need to be in any special order. When the blocking column is text, the first entry (row 1) is block 1, the next entry that is different is block 2, and so on. When the blocking column is numeric, the smallest value is block 1, the second smallest value is block 2, and so on. If you use BLOCK without REPS, Minitab stores a column of 1s.

RBLOCK

Specifies blocks on replicates. Each set of replicate points are placed in a separate block. No arguments are taken.

LEVELS K K,..., K K

LEVELS C

Specifies factor levels for each factor.

Use LEVELS to specify values for the levels of each factor. The level values can be text or numeric. If you do not use the subcommand, the levels are the integers 1, 2, 3, and so on.

With LEVELS K K, ..., K K, you can set levels to be text or numeric data. The arguments that define text data must be enclosed in double quotation marks. List a pair of arguments for each factor. The first K is the low level. The second K is the high level.

With LEVELS C, the level values must be numeric. Put the level values in a column, in the same order as if they were listed on the subcommand. Then, specify the level column on the LEVELS subcommand.

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

The optional K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output for the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

FFDESIGN: Session command for creating a full or fractional factorial design

FFDESIGN K [K]

Creates a full or fractional factorial design with K factors and K number of runs. These designs are often referred to as screening designs. If the second K is omitted and the first K is 7 or less, a full factorial design is generated. To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, CTPT, and XMATRIX.)

Minitab has a catalog of designs. For a summary of the catalog, go to Designs generated by FFDESIGN on page 1224. The two constants (K) on FFDESIGN specify a design from that catalog. The first K is the number of factors, and the second K is the number of runs. You can then use REPS to replicate points, FOLD the design, and add center points (CPBLOCK). You may also ADD additional factors, provided the total number is at most 15.

The design generators, resolution, and alias structure are displayed. The design matrix is displayed under BRIEF 3. The defining relation is displayed under BRIEF 4. Factors are denoted by the letters A, B, C, D, E, F, G, H, J, K, L, M, N, O, P on the output and in the subcommands BLOCKS, ADD, and FOLD.

Note The letter I is not used.

The alias structure that is displayed by default depends on the number of factors. You can use ALIAS to change what is displayed.

Number of factors	What is displayed
2–6	All terms
7–10	Up to three-way interactions

Number of factors	What is displayed
11–15	Up to two-way interactions

Quotation marks around column names are required. If you use a subcommand more than one time, FFDESIGN uses the last instance of the subcommand.

REPS K

Replicates each point K times. Specifies up to 50 replicates of each point. REPS does not replicate center points.

The following example generates a full factorial design that has 4 factors and 2 observations per cell. Thus, there are a total of 32 runs in this experiment.

```
FFDESIGN 4 16;
  REPS 2.
```

BLOCKS

Specifies the number of blocks in K, or specifies the terms to use as block generators in termlist.

If you use the subcommand ADD, you must also specify your own block generators. You cannot use an integer with BLOCKS when you specify ADD. Minitab cannot determine good block generators in this case. No extra text is allowed on the BLOCKS line, except after a #.

Blocking a design can reduce its resolution. Let r_1 = the resolution before blocking. Let r_2 = the length of the shortest term that is confounded with blocks. Then the resolution after blocking is the smaller of r_1 and (r_2 + 1).

BLOCKS K

With BLOCKS K, you specify the number of blocks that you want, from 2 to 99. This number must evenly divide the total number of runs, excluding center points (those given on FFDESIGN times those given on REPS). Minitab then chooses an appropriate design.

If your design has replicates and you specify an integer on BLOCKS, then Minitab attempts to put replicates in different blocks. The exact rule is a bit complicated.

BLOCKS *termlist*

With BLOCKS termlist, you specify terms to use as block generators.

The following example generates a 64 run design with 8 factors.

```
FFDESIGN 8 64;
 BLOCKS ABC CDE.
```

The factors are labeled A, B, C, D, E, F, G, H. The block generators are ABC and CDE. This gives four blocks. The blocks are numbered in standard order (also called Yates order). The blocks are as follows.

Block	ABC	CDE
1	–	–
2	+	–
3	–	+
4	+	+

CPBLOCK K

Adds center points to the design.

Note CPBLOCK supersedes the CENTER subcommand.

CPBLOCK adds center points as follows:

- When all factors are numeric, Minitab adds K center points as follows:
  - If the design is not blocked, then Minitab adds K center points to the design.
  - If the design is blocked, then Minitab adds K center points to each block.
- When all factors are text, you cannot use CPBLOCK to add center points.
- When factors are both numeric and text, the design does not have a true center. In this case, center points are called pseudo-center points, and Minitab adds K center points as follows:
  - If the design is not blocked, then Minitab adds K center points for each combination of the levels of the text factors.
  - If the design is blocked, then Minitab adds K center points for each combination of the levels of the text factors to each block.

For example, consider an unblocked 2**3 design, factors A and C are numeric. Factor B is text that indicates whether a catalyst is absent or present. The LEVELS subcommand assigns text data for the levels of factor B.

```
FFDESIGN 3 8;
CPBLOCK 3;
BLOCKS 1;
LEVELS 0 10 "absent" "present" .2 .3
DESIGN.
```

Factor B has 2 levels: present or absent. With `CPBLOCK 3`, Minitab adds a total of 2 x 3 = 6 center points, three points for the low level of factor B and 3 for the high level. The six points are as follows:

5 present .25

5 present .25

5 present .25

5 absent .25

5 absent .25

5 absent .25

Next, consider a blocked 2**5 design where 3 factors are text. There are 2 x 2 x 2 = 8 combinations of text levels. If you use `CPBLOCK 2` and `BLOCKS 2`, then Minitab adds 8 x 2 = 16 pseudo-center points to each of the 2 blocks.

LEVELS

Use LEVELS to specify values for the levels of each factor. The level values can be text or numeric. The default values are –1 for the low level, +1 for the high level, and 0 for a center point.

LEVELS K K, ..., K K

With LEVELS K K, ..., K K, you list one pair of arguments for each factor. The first K should be the low level, and the second K should be the high level. When you add center points with CPBLOCK, Minitab calculates the center points as the value that is half way between the low and high levels.

The following example stores 0 for the low level of A, 10 for the high level, and 5 for the center point; 100 for the low level of B, 102 for the high level, and 101 for the center point; .2 for the low level of C, .3 for the high level, and 0.25 for the center point.

```
FFDESIGN 3 8;
 CPBLOCK 2;
 LEVELS  0 10 100 102 .2 .3;
 BLOCKS 1;
 DESIGN
```

With LEVELS K K, ..., K K, you can set levels to be text data. The arguments that define text data must be enclosed in double quotation marks. For example, if factor B is whether a catalyst is present or absent, you can use the following:

```
FFDESIGN 3 8;
  CPBLOCK 2;
  LEVELS  0 10 "absent" "present" .2 .3;
  BLOCKS 1;
  DESIGN.
```

Because you use `CPBLOCK 2` to add center points with a text factor, Minitab creates pseudo-center points. For more information, see the CPBLOCK subcommand.

LEVELS C

With LEVELS C, the level values must be numeric. Put the level values in a column, in the same order as if they were listed on the subcommand. Then, specify the level column on the LEVELS subcommand.

FOLD ***[factor...factor]***

Folds the design on the specified factors. If you don't include factors, Minitab folds the design on all factors.

The signs on the factors that you specify are switched. The signs on the remaining factors are not switched. These rows for the non-specified factors are then appended to the end of the design matrix, which doubles the number of runs.

Consider the following example.

```
FFDESIGN 3 4;
   FOLD A.
```

First, Minitab generates a 3 factor design in 4 runs. The first 4 rows of the data matrix below are these runs. Then, Minitab folds on factor A. This adds 4 additional runs.

A	B	C
–	–	+
+	–	–
–	+	–
+	+	+
+	–	+
–	–	–
+	+	–
–	+	+

Folding is a way to reduce confounding. If you fold on one factor, say A, then A and all its two-factor interactions are free from other main effects and two-factor interactions. If you fold on all factors, then all main effects are free from all two-factor interactions.

When you fold, the defining relation is usually shortened. Specifically, any word in the defining relation that has an odd number of the letters on which you folded the design is omitted. If you fold a design and the defining relation is not shortened, then the folding adds replicates, but it does not reduce confounding. In this case, Minitab displays an error message.

If you fold a design that is blocked, the same block generators are used for the folded design as for the unfolded design.

FRACTION K

Specifies which fraction of a fractional design to use. The default fraction is called the principal fraction. This is the fraction where all signs are positive. Consider the following example.

```
FFDESIGN 5 8;
  FRACTION 3.
```

A full design with 5 factors requires 32 runs. With just 8 runs, a one-fourth fraction is used. You can use any of the four possible fractions of the design. Minitab numbers them in standard order (also called Yates' order) using the design generators, as follows.

1	D = –AB	E = –AC
2	D = AB	E = –AC
3	D = –AB	E = AC
4	D = AB	E = AC

The default fraction is called the principal fraction. This is the fraction where all signs are positive. `FRACTION 3` specifies the third fraction. This is the one with design generators D = –AB, E = AC.

FRACTION applies to all the factors: the ones specified on the command line and any you added with ADD. If you use both FRACTION and ADD, you cannot use minus signs in the design generators that are listed on ADD. Actually, in this case, minus signs on ADD are not useful.

Sometimes a design contains points that are impractical to run. Choosing an appropriate fraction can avoid these points. For example, suppose you could not run the design in the example with all five factors set at their high level. The principal fraction contains this point, but the third fraction does not.

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block, and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any positive integer as an argument. To always get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, center point indicator, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output.

You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays a summary of the design.
2 (default)	In addition to the K = 1 output, Minitab displays the alias table.
3	In addition to the K = 2 output, Minitab displays the design table.
4	In addition to the K = 3 output, Minitab displays the defining relation. If the design is a full factorial, there is no defining relation, so K = 4 is equivalent to K = 3.

ALIAS K

Specifies the highest-order interaction to print in the alias table. For example, `ALIAS 3` displays single factors, and two- and three-way interactions.

Important A large K with a large number of factors might take a very long time to calculate.

ADD *factor = term...factor = term*

ADD is used primarily when you want to create custom designs rather than use a design in Minitab's catalog. With ADD, you can add factors to a design by specifying the appropriate generators.

Consider the following example.

```
FFDESIGN 3 8;
 ADD D=AB E=AC.
```

The example starts with three factors labeled A, B, and C. Then, factors D and E are added. D is confounded with the AB interaction. E is confounded with the AC interaction. This gives a $2^{(5-2)}$ design. You can also use a minus interaction (for example, D = –AB). The added factors must be given in alphabetical order. The total number of factors, those listed on FFDESIGN plus those listed on ADD, can be at most 15.

No extra text is allowed on the ADD line, except after a #.

If you use BLOCKS with ADD, you must specify your own block generators. You cannot use an integer with BLOCKS when you specify ADD. Minitab cannot determine good block generators for a user-specified design.

SPDESIGN: Session command for creating a 2-level split-plot design

SPDESIGN K K

Creates 2-level split-plot designs, either full or fractional. In a split-plot experiment, levels of the hard-to-change factor are held constant for several experimental runs. The experimental runs are collectively treated as a whole plot. The easy-to-change factors are varied over these runs. Each combination of the factors is considered a subplot. To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.)

You can have 2-7 factors in a split-plot experiment. The number of hard-to-change factors that you can have depends on the number of factors. For 2 factors, you can have only 1 hard-to-change factor. For 3 factors, you can have 1 or 2 hard-to-change factors. For 4, 5, 6, or 7 factors, you can have 1, 2, or 3 hard-to-change factors.

You specify a full or fractional design with the PLOTS subcommand.

PLOTS K K

Specifies K whole plots with K subplots.

WREPLICATES K

Replicates the base design by multiplying the number of whole plots by K.

SREPLICATES K

Replicates the subplots within each whole plot K times.

RBLOCK

Specifies blocks on whole block replicates. Each replicate of the base design is placed in a separate block. No arguments are taken.

LEVELS

Specifies the values for the levels of each factor. The level values can be text or numeric. The default values are –1 for the low level and +1 for the high level.

The two ways to use LEVELS are as follows.

LEVELS K K, ..., K K

With LEVELS K K, ..., K K, you can set levels to be text or numeric data. The arguments that define text data must be enclosed in double quotation marks. List a pair of arguments for each factor. The first K should be the low level. The second K should be the high level.

In the following example, factor B is whether a catalyst is present or absent.

```
SPDESIGN 2 1;
  PLOT 2 2;
  WPREPLICATES 2;
  RANDOMIZE;
  ALIAS 2;
  LEVELS  0 10 "absent" "present";
  DESIGN.
```

LEVELS C

With LEVELS C, the level values must be all numeric or all text. First, put the level values in a column, in the same order as if they were listed on the subcommand. Then, specify this column on the LEVELS subcommand.

FRACTION K

Specifies which fraction of a fractional design to use with K. K must be a positive integer.

A full design that has 5 factors requires 32 runs. A full factorial design has no fractions, so no value of K is legal. The legal values for a ½-fractional design are 1 and 2. The legal values for a ¼-fraction are 1, 2, 3, and 4. The following example is a design with 2 whole plots and 8 subplots, which makes a ½-fraction.

```
SPDESIGN 5 1;
  PLOT 2 8;
  FRACTION 2.
```

As in FFDESIGN, the fractions are numbered according to the signs that are used in the defining relation.

Fraction number	Defining relation
1	I – ABCD - CDE + ABE
2	I + ABCD – CDE – ABE
3	I – ABCD + CDE – ABE
4	I + ABCD + CDE + ABE

The default fraction is called the principal fraction. This is the fraction for which all signs are positive in the defining relation.

Sometimes a design contains points that are impractical to run. Choose an appropriate fraction to avoid these points.

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. BRIEF can also be used as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that Minitab displays
0	Minitab does not display output, but does complete all specified storage.
1	Minitab displays a summary of the design.
2 (default)	In addition to the K = 1 output, Minitab displays the alias table.
3	In addition to the K = 2 output, Minitab displays the design table.
4	In addition to the K = 3 output, Minitab displays the defining relation. If the design is a full factorial, there is no defining relation, so K = 4 is equivalent to K = 3.

ALIAS K

Specifies the highest order interaction to display in the alias table. For example, ALIAS 3 displays single factors, and 2- and 3-way interactions.

Important A large K along with a large number of factors might take a very long time to calculate.

PBDESIGN: Session command for creating a Plackett-Burman design

PBDESIGN K [K]

Creates Plackett-Burman designs. Plackett-Burman designs are a class of resolution III, 2-level fractional factorial designs that are often used to study main effects. In a resolution III design, main effects are aliased with 2-way interactions.

The two constants on PBDESIGN specify a design. The first K is the number of factors. The second K is the number of runs. Factor columns can be text or numeric.

To see the designs that Minitab generates, go to Plackett-Burman designs on page 1259. A design is based on the number of runs, from 8 to 48 and always a multiple of 4. Therefore, the second K on PBDESIGN must be a multiple of 4. The number of factors must be less than the number of runs. If you don't include the second K, Minitab sets the number of runs to the smallest possible value for the specified number of factors.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTPYE, and XMATRIX.)

CENTER K

Adds up to 50 center points to the design. If the design is blocked, the center points are divided equally between the blocks. If K is not a multiple of the number of blocks, then each of the last few blocks has one less point than the other blocks.

REPS K

Specifies up to 50 replicates of each point. REPS does not replicate center points.

The following example generates a Plackett-Burman design that has 2 factors and 8 base runs. Thus, there are a total of 16 runs in this experiment.

```
PBDESIGN 2 8;
 REPS 2.
```

RBLOCK

Specifies blocks on replicates. Each set of replicate points are placed in a separate block. No arguments are taken.

LEVELS

Use LEVELS to specify values for the levels of each factor. The level values can be text or numeric. The default values are -1 for the low level, +1 for the high level, and 0 for a center point.

The two ways to use LEVELS are as follows.

LEVELS K K, ..., K K

With LEVELS K K, ..., K K, you list one pair of arguments for each factor. The first K should be the low level and the second K should be the high level. When you add center points with CPBLOCK, Minitab calculates the center points as the value that is half way between the low and high levels. The following example stores 0 for the low level of A, 10 for the high level, and 5 for the center point. 100 for the low level of B, 102 for the high level, and 101 for the center point; .2 for the low level of C, .3 for the high level, and 2.5 for the center point.

```
PBDESIGN 2 8;
  CENTER 2;
 LEVELS  0 10 100 102;
 DESIGN.
```

With LEVELS K K, ..., K K, you can set levels to be text data. The arguments that define text data must be enclosed in double quotation marks. In the following example, factor B is whether a catalyst is present or absent. Because you use CENTER 2 to add center points with a text factor, Minitab creates pseudo center points.

```
PBDESIGN 2 8;
  CENTER 2;
  RBLOCK;
  LEVELS  0 10 "absent" "present";
  DESIGN.
```

LEVELS C

With LEVELS C, the level values must be numeric. Put the level values in a column, in the same order as if they were listed on the subcommand. Then, specify the level column on the LEVELS subcommand.

RANDOMIZE [K]

Randomizes the runs in the design matrix. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output for the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

FFACTORIAL: Session command for analyzing a full or fractional factorial design

Analyzes a factorial design with a continuous response. Fits full and fractional factorial models, including those produced by FFDESIGN on page 280 and PBDESIGN on page 259. Each factor must have two levels. Also, you can have covariates and a blocking factor. The design does not have to be orthogonal. You can specify a single response. Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When you include center points in the model, Minitab tests for curvature.

You can display effects plots and residual plots. You can store the residuals, fitted values, and many other diagnostics for further analysis. You can use extra text on the command and subcommand lines only after a #. Quotation marks around column names that are one word long are optional.

For information about how Minitab calculates the model, go to Calculations for FFACTORIAL on page 1219.

For information about how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

RESPONSE C

Specifies the column that contains the response variable in C. The column must be numeric or date/time.

FACTOR C...C (required)

Specifies the factors in columns C...C. The columns must have the same number of rows that the response column has.

COVARIATES C...C

Specifies columns that include covariates in C...C. Include the covariate columns with terms to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

BLOCKS C (optional)

Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

CTPT C

Specifies the column that identifies which design points are center points in C. Include the center point column together with TERMS to make a model that tests for curvature.

TERMS *termlist*

Specifies the terms to include in the model. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

LEVELS K K,..., K K

LEVELS C

Specifies the uncoded factor levels for each factor. In a pair of levels, the first K is the low level and the second K is the high level. In a pair of column rows, the first row is the low level and the second row is the high level. For each factor, the list or column must contain two values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all of the levels are continuous or all of the levels are categorical.

Options

WEIGHTS C

Specifies a numeric column (C) that contains weights to use in weighted regression analysis. The same weights are used for all responses. If different responses require different weights, then FFACTORIAL will need to be run as many times as there are different weights.

Typically, the weights are the reciprocal of the variances of the responses. In this case, weights are based on the fitted variances calculated and stored by VFACTORIAL. You can use other weight variables.

The leverages stored by HI are affected by weights.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

BOXCOX [K]

Transforms each variable or group to follow a normal distribution using the lambda value in K. K must be between –5 and +5. If you do not specify K, then Minitab finds the optimal lambda. By default, Minitab rounds the optimal value.

To use the exact lambda, change the setting in **File** > **Options** > **Linear Models** > **Display of Results**.

Minitab cannot calculate the optimal lambda for stepwise regression. So you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

MEANS *termlist*

Calculates the least squares means for the terms that you specify in the termlist. You can use the following commands to display and store the least squares means.

TMEANS

Displays the least squares means for the specified terms.

SMEANS

Stores the least squares means for the specified terms.

BSMEANS

If you transformed the data, stores the least squares means for the specified terms.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors. By default, K = 1E-12.

The tolerance value determines when Minitab retains a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 – R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 – R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

Stepwise regression

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE,

unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]

Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C

Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

The following subcommands specify what is displayed in the results.

NODEFAULT

Specifies that no default tables or graphs are displayed.

TEXPAND

Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE

Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TMETHOD

Displays the method table.

TMSDETAILS

Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL

Displays the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

NOFULL

Does not display the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

TEQUATION

Displays the regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TCOEFFICIENT

Displays the table of coefficients. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

Displays the summary of model table, which includes the R^2 statistics and S. The expanded version of the table includes the prediction sum of squares (PRESS), the corrected Akaike's Information Criterion (AICc) and the Bayesian Information Criterion (BIC).

TANOVA

Displays the ANOVA table. The expanded version of the table includes the sequential sums of squares and the percent contribution to the total sum of squares.

TDIAGNOSTICS [K]

Displays the table of diagnostics. The expanded version of the table includes the standard error of the fit, the confidence interval for the fit, the deleted residual, the leverage, Cook's D, and DFITS.

Value of K	Diagnostics that are displayed
0	Only for unusual observations
1	For all observations

TALIAS [K]

Displays the alias table up to order K. One line is displayed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, then Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to three-way interaction for 7 to 10 factors, and two-way interactions for more than 10.

TMEANS

Displays the table of least squares means for all combinations of the factor levels that you specify in MEANS.

Storage

The following subcommands specify what is stored in the worksheet.

FITS C

Stores the fitted values.

RESIDUALS C

Stores the residuals (response – fit) in C.

SRESIDUALS C

Stores the standardized residuals.

TRESIDUALS C

Stores the deleted Studentized residuals in C.

HI C

Stores the leverage values.

COOK C

Stores Cook's distance values, using one column for each response.

DFITS C

Stores DFITS (also called DFFITS), using one column for each response.

EFFECTS C

Stores the effects that are included in the output in C.

Note Effects are not displayed or stored for the constant, covariates, or blocks.

COEFFICIENTS C

Stores the coefficients that are included in the coefficients table in C.

The coefficients are based on coded levels (–1, 0, +1), not on the actual levels in the data set. A coefficient is always half the value of the corresponding effect.

XMATRIX *M*

Stores the design matrix. The first column is a column of ones for the constant term. Next, there is one column for each covariate. Next, if the design was blocked into k blocks, there are (k – 1) columns. Next, there is one column for each factor, using –1 for the low level, 0 for a center point, and +1 for the high level. Next, there is

one column for each interaction in the model. The column for an interaction is the product of all the columns that correspond to the factors that are contained in that interaction.

If FFACTORIAL or VFACTORIAL removed some of the interactions that you specified in your model, the interactions are not included in the stored design matrix. The columns in XMATRIX match the coefficients that are displayed and stored.

BCRESP C

Stores the Box-Cox transformation of the response.

BFITS C

Stores the back-transformed fits after the Box-Cox transformation in C.

BSMEANS C

If you use a Box-Cox transformation, BSMEANS stores the means for the back-transformed response in C.

SMEANS C

Stores the least squares means from MEANS in C.

If you use a Box-Cox transformation, SMEANS stores the means for the transformed response in C.

BFFACTORIAL: Session command for analyzing a full or fractional factorial design with a binary response

Required commands

To perform the analysis, use the main command and at least these accompanying subcommands.

BFFACTORIAL

Analyzes a factorial design with a binary response. Fits full and fractional factorial models, including those produced by FFDESIGN on page 280 and PBDESIGN on page 259. Each factor must have two levels. Also, you can have covariates and a blocking factor. The design does not have to be orthogonal. You can specify a single response. Specify interactions in the following way: C5*C6. When you include center points in the model, Minitab tests for curvature. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

This analysis provides diagnostic plots, goodness-of-fit tests, and other diagnostic measures so you can assess the validity of your model. For more information, go to Generalized linear model diagnostics and residual analysis on page 1260.

Minitab excludes all observations with missing values on either the response variable or any of the predictors from all calculations.

RESPONSE C C

Enters the response data as a column that contains the number of events first and then a column that contains the number of trials. The columns must have the same number of rows.For examples, go to Entering data for response variables on page 1225.

FACTOR C...C

Specifies the factors in columns C...C. The columns must have the same number of rows as the response columns.

TERMS *termlist*

Specifies the terms to include in the model. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

LEVELS K K,..., K K or LEVELS C

Specifies the uncoded factor levels for each factor. In a pair of levels, the first K is the low level and the second K is the high level. In a pair of column rows, the first row is the low level and the second row is the high level. For each factor, the list or column must contain two values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all of the levels are continuous or all of the levels are categorical.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

Optional subcommands

Use these subcommands and the subcommands in the following sections to change the default analysis.

EVNAME K

Specify the name of the event with K. Enclose text values in double quotation marks.

LOGIT

Use the logit link function.

NORMIT

Use the normit link function.

GOMPIT

Use the gompit link function. The gompit link function is also called the complementary log-log link function.

BLOCKS C

Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

COVARIATES C...C

Specifies columns that include covariates in C...C. Include the covariate columns with TERMS to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

CTPT C

Specifies the column that identifies which design points are center points in C. Include the center point column together with TERMS to make a model that tests for curvature.

Options subcommands

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SEQDEVIANCE

Use sequential deviances for the tests of significance in the ANOVA table. The order of the terms in TERMS affects the statistical significance. The default is to use adjusted deviances.

PEARSON

Specifies the use of Pearson residuals. The default is to use deviance residuals.

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} (\mathbf{X}' \mathbf{W} \mathbf{Y})$

HGROUP K

Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test with K. The default is 10.

MEANS *termlist*

Calculates the fitted mean probability for the terms that you specify in the termlist. You can use the following commands to display and store the probabilities.

TMEANS

Displays the fitted mean probability for the specified terms.

SMEANS C

Stores the fitted mean probability for the specified terms.

ITERATION K

Specifies the maximum number of iterations for the optimization algorithm that estimates the coefficients. The default is 50.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors. By default, K = 1E-12.

The tolerance value determines when Minitab retains a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

CTOLERANCE K

Specifies the convergence criterion for the algorithm that estimates the coefficients. By default, K = 1E-8.

START C

Specifies a column that contains starting values for the coefficients in the optimization algorithm. The column begins with the estimate of the constant, then follows the order of the model after TERMS.

Stepwise regression subcommands

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs subcommands

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS
Displays residuals versus fitted values.

GORDER [C]
Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]
Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C
Displays the residuals versus the variables in C. Typically, you enter the columns for the predictors.

Results subcommands

NODEFAULT
Hides default tables and graphs that you do not include the subcommand to display.

TEXPAND
Displays the expanded versions of the tables that have them. The expanded Analysis of Variance table adds the sequential deviances and percent contributions for each term. The Coefficients table adds confidence intervals for the coefficients. The Fits and Diagnostics table adds the standard error of the fit, the 95% confidence interval for the fit, the deleted residual, the HI, the Cook's D, and the DFITS.

TSIMPLE
Displays the simple versions of the Analysis of Variance table, the Coefficients table, and the Fits and Diagnostics table.

TMETHOD
Displays a table of information about the method and options.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model.

If you do not specify FULL or NOFULL, then Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Specifies to display the coefficients, p-values, and model summary statistics for each step of the procedure.

NOFULL
Hides these statistics.

TRINFO
Displays the number of events, non-events, and the total. The table also indicates which response value is the reference event.

TSTEP
Displays the deviance at each iteration of the coefficients.

TCOEFFICIENTS
Displays the coefficients and the p-values for the Wald normal approximation tests.

If you do not specify FULL or NOFULL, Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL
Specifies to show the coefficients for the reference level when you use (0, 1) coding for categorical predictors.

NOFULL
Hides these statistics.

TODDS
Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model.

INCREMENT K...K
Specifies the increment for the odds ratio for the predictor with K. For example, if the predictor is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram. Enter a value for each numeric factor in the same order that the factors follow the FACTOR subcommand. Then, enter a value for each covariate in the order that the covariates follow the COVARIATE subcommand. Enter 1 for a factor or covariate to use the same units as the data.

TSUMMARY
Displays the statistics that evaluate model fit, including deviance R^2.

TGOODNESS
Displays the Pearson and deviance goodness-of-fit tests. For the binomial distribution models, also shows the Hosmer-Lemeshow goodness-of-fit test.

TDEVIANCE
Displays the ANOVA table.

TEQUATION
Displays the regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TMEANS
Displays the table of fitted mean probabilities for all combinations of the factor levels that you specify in MEANS.

THOSMER
Displays the table of observed and expected values for the Hosmer-Lemeshow test.

TASSOCIATION
Displays the number of concordant and discordant pairs, the Somers' D, the Goodman-Kruskal Gamma, and Kendall's Tau-a.

TDIAGNOSTICS [K]
Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

TALIAS [K]
Displays the alias table up to order K. One line is displayed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, then Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to three-way interaction for 7 to 10 factors, and two-way interactions for more than 10.

Storage subcommands

FITS C
Stores the event probabilities in C.

RESIDUALS C
Stores the deviance residuals by default in C. Stores the Pearson residuals if you use the subcommand PEARSON.

SRESIDUALS C
Stores the standardized deviance residuals by default in C. Stores the standardized Pearson residuals if you use the subcommand PEARSON.

TRESIDUALS C
Stores the deleted deviance residuals by default in C. Stores the deleted Pearson residuals if you use the subcommand PEARSON.

COOK C
Stores the Cook's distances in C.

DFITS C
Stores the DFITS in C.

HI C
Stores the leverages in C.

SADJUSTEDRESP C
Stores the final adjusted response values in C.

SIWEIGHTS C
Stores the internal weights used to estimate the model parameters in C.

SECOEFFICIENTS C
Stores the standard errors of the estimated coefficients in C.

COEFFICIENTS C
Stores the estimated coefficients in C.

EFFECTS C
Stores the effects that are included in the output in C.

Note Effects are not displayed or stored for the constant, covariates, or blocks.

SMEANS C
Stores the fitted mean probability for the specified terms in C.

DBETA C
Stores the change in the estimated regression coefficients (delta b) when you delete all observations with a particular factor/covariate pattern in C. Use the delta b values to detect observations with a strong influence on the coefficients.

DSBETA C
Stores the change in the estimated regression coefficients when you delete all observations with a particular factor/covariate pattern based on the standardized Pearson residual in C.

DCHISQUARE C

Stores the change in the chi-square statistic when you delete all observations with a particular factor/covariate pattern in C. Observations that are poorly fit by the model have high delta chi-square values.

DDEVIANCE C

Stores the change in the deviance when you delete all observations with a particular factor/covariate pattern in C.

XMATRIX *M*

Stores the design matrix in M.

XPWXINVERSE *M*

Stores the X'WX inverse matrix of the estimated coefficients in M.

GFACTORIAL: Session command for fitting a general full factorial design

GFACTORIAL

Analyzes a general factorial design that can have categorical predictors and continuous covariates using least squares regression.

Fits general full factorial models, including those produced by FDESIGN on page 279. In addition, you can have covariates and a blocking factor.

For general full factorial designs, you specify the number of factors and the levels of each factor. You can have arbitrary levels but no fractional designs.

The design does not have to be orthogonal. You can specify up to 25 responses and have 2 to 15 factors. Each factor must have at least 2 levels and at most 100 levels. The number of levels can be different for each factor.

Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

Optionally, you can store the residuals, fitted values, and many other diagnostics for further analysis.

You can use extra text on the command and subcommand lines only after a #. Quotation marks around column names are optional.

For information about how Minitab calculates the model, go to Calculations for FFACTORIAL on page 1219.

For information on how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

BLOCKS C (optional)

Specifies the column that contains the blocking information. Include the blocking column together with TERMS to make a model that considers blocks.

FACTOR C...C

Specifies the categorical factors, if you have any. The columns can be numeric, text, or date/time. The columns must have the same number of rows as the response column.

NLEVELS K...K

Use both NLEVELS and LEVELS when the factor levels in the worksheet are coded and you want to display uncoded levels in the results.

NLEVELS specifies the number of levels for each factor. Each K must be an integer from 2 to 100. You must specify the same number of K arguments as the number of columns that you specify in FACTOR.

LEVELS *E...E*

Use both NLEVELS and LEVELS when the factor levels in the worksheet are coded and you want to display uncoded levels in the results.

LEVELS specifies the uncoded values for all the factor levels. You must specify the same number of values for LEVELS as the number of arguments that you specify for NLEVELS. The group of values for each factor must be of the same type.

COVARIATES C...C

Specifies which columns of variables are continuous. Include the covariate columns with terms to make a model that considers the covariates. Covariates are fit first, then the blocks, then all other terms. You can have up to 50 covariates.

TERMS *termlist*

Specifies the terms to include in the model. Terms can contain a maximum of two factors. Only continuous predictors can be repeated. Nested terms are not allowed. If terms are to be fitted, they must contain the blocks. The model can be nonhierarchical.

Options

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by the following:

$(\mathbf{X'\,W\,X})^{-1}\,(\mathbf{X'\,W\,Y})$

This is equivalent to minimizing the weighted SS Error as follows:

$\Sigma\, w_i\,(Y -)^2$ where w_i is the weight in row i

The prediction interval that is calculated by PREDICT assumes a weight of 1. If you use the WEIGHT subcommand with values other than 1, you should adjust the prediction interval values manually.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

TOLER K

Specifies the tolerance level for collinearity and constant check. By default, K = 1E–12.

Box-Cox

BOXCOX [K]

Transforms each variable or group to follow a normal distribution using the lambda value in K. K must be between –5 and +5. If you do not specify K, then Minitab finds the optimal lambda. By default, Minitab rounds the optimal value.

To use the exact lambda, change the setting in **File** > **Options** > **Linear Models** > **Display of Results**.

Minitab cannot calculate the optimal lambda for stepwise regression. So you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

Means

MEANS *termlist*

Calculates least squares means for the specified terms. You can use the following commands to display and store the least squares means.

TMEANS

Displays the calculated least squares means for the specified terms.

SMEANS

Stores the calculated least squares means for the specified terms.

BSMEANS

If you transformed the data, stores the calculated least squares means for the specified terms.

Stepwise

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1

degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS
Specifies that both categorical and continuous terms have to be hierarchical.

ATEND
Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS
Specifies that the model is hierarchical at every step.

SINGLE
Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE
Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

GPARETO
Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K
Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL
Specifies that only the terms that are in the model are on the plots.

RTYPE K
Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER [C]
Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]
Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C
Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

The following subcommands specify what is displayed in the results.

NODEFAULT
Specifies that no default tables or graphs are displayed.

TEXPAND
Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE
Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TMETHOD
Displays the method table.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Displays the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

NOFULL
Does not display the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

TFACTOR
Displays information about the factors.

TEQUATION
Displays the regression equation table.

TCOEFFICIENT

Displays the table of coefficients. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

Displays the summary of model table, which includes the R^2 statistics and S. The expanded version of the table includes the prediction sum of squares (PRESS), the corrected Akaike's Information Criterion (AICc) and the Bayesian Information Criterion (BIC).

TANOVA

Displays the ANOVA table. The expanded version of the table includes the sequential sums of squares and the percent contribution to the total sum of squares.

TDIAGNOSTICS [K]

Displays a table of diagnostics. The expanded version of the table includes the standard error of the fit, the confidence interval for the fit, the deleted residual, the leverage, the Cook's D, and the DFITS.

Value of K	Diagnostics that are displayed
0	Diagnostics only for unusual observations
1	Diagnostics for all observations

TMEANS

Displays the least squares means for all combinations of the factor levels that you specify in MEANS.

Storage

The following subcommands specify what is stored in the worksheet.

BCRESP C

Stores the Box-Cox transformation of the response.

BFITS C

Stores the fits for the original response.

FITS C

Stores the fitted values.

SMEANS C

Stores the least squares means for all combinations of factor levels that you specified in MEANS, in C.

BSMEANS C

Stores the least squares means for the original response, if the data were transformed. Stores all combinations of factor levels that you specified in MEANS, in C.

RESIDUALS C

Stores the residuals (fitted values – observed values).

SRESIDUALS C
Stores the standardized residuals.

TRESIDUALS C
Stores the deleted Studentized residuals in C.

HI C
Stores the leverage values.

COOK C
Stores the Cook's distance values, using one column for each response, in C.

DFITS C
Stores the DFITS (also called DFFITS), using one column for each response, in C.

COEFFICIENTS C
Stores the estimated coefficients.

XMATRIX *M*
Stores the design matrix for the regression model in M.

BGFACTORIAL: Session command for analyzing a general full factorial design with a binary response.

Required commands

To perform the analysis, use the main command and at least these accompanying subcommands.

BGFACTORIAL

Analyzes a general factorial design that can have categorical predictors and continuous covariates with a binary response.

Fits general full factorial models, including those produced by FDESIGN on page 279. In addition, you can have covariates and a blocking factor.

For general full factorial designs, you specify the number of factors and the levels of each factor. You can have arbitrary levels but no fractional designs.

The design does not have to be orthogonal. You can specify up to 25 responses and have 2 to 15 factors. Each factor must have at least 2 levels and at most 100 levels. The number of levels can be different for each factor.

Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for factorial designs on page 1249.

This analysis provides diagnostic plots, goodness-of-fit tests, and other diagnostic measures so you can assess the validity of your model. For more information, go to Generalized linear model diagnostics and residual analysis on page 1260.

Minitab excludes all observations with missing values on either the response variable or any of the predictors from all calculations.

RESPONSE C C

Enters the response data as a column that contains the number of events first and then a column that contains the number of trials. The columns must have the same number of rows.For examples, go to Entering data for response variables on page 1225.

FACTOR C...C

Specifies the factors in columns C...C. The columns must have the same number of rows as the response columns.

TERMS *termlist*

Specifies the terms to include in the model. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

Optional subcommands

Use these subcommands and the subcommands in the following sections to change the default analysis.

EVNAME K

Specify the name of the event with K. Enclose text values in double quotation marks.

LOGIT

Use the logit link function.

NORMIT

Use the normit link function.

GOMPIT

Use the gompit link function. The gompit link function is also called the complementary log-log link function.

BLOCKS C

Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

COVARIATES C...C

Specifies columns that include covariates in C...C. Include the covariate columns with TERMS to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

NLEVELS K...K

Use both NLEVELS and LEVELS when the factor levels in the worksheet are coded and you want to display uncoded levels in the results.

NLEVELS specifies the number of levels for each factor. Each K must be an integer from 2 to 100. You must specify the same number of K arguments as the number of columns that you specify in FACTOR.

LEVELS *E...E*

Use both NLEVELS and LEVELS when the factor levels in the worksheet are coded and you want to display uncoded levels in the results.

LEVELS specifies the uncoded values for all the factor levels. You must specify the same number of values for LEVELS as the number of arguments that you specify for NLEVELS. The group of values for each factor must be of the same type.

Options subcommands

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SEQDEVIANCE

Use sequential deviances for the tests of significance in the ANOVA table. The order of the terms in TERMS affects the statistical significance. The default is to use adjusted deviances.

PEARSON

Specifies the use of Pearson residuals. The default is to use deviance residuals.

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} (\mathbf{X}' \mathbf{W} \mathbf{Y})$

HGROUP K

Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test with K. The default is 10.

MEANS *termlist*

Calculates the fitted mean probability for the terms that you specify in the termlist. You can use the following commands to display and store the probabilities.

TMEANS

Displays the fitted mean probability for the specified terms.

SMEANS C

Stores the fitted mean probability for the specified terms.

ITERATION K

Specifies the maximum number of iterations for the optimization algorithm that estimates the coefficients. The default is 50.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors. By default, K = 1E-12.

The tolerance value determines when Minitab retains a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

CTOLERANCE K

Specifies the convergence criterion for the algorithm that estimates the coefficients. By default, K = 1E-8.

START C

Specifies a column that contains starting values for the coefficients in the optimization algorithm. The column begins with the estimate of the constant, then follows the order of the model after TERMS.

Stepwise regression subcommands

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs subcommands

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS
Displays residuals versus fitted values.

GORDER [C]
Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]
Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C
Displays the residuals versus the variables in C. Typically, you enter the columns for the predictors.

Results subcommands

NODEFAULT
Hides default tables and graphs that you do not include the subcommand to display.

TEXPAND
Displays the expanded versions of the tables that have them. The expanded Analysis of Variance table adds the sequential deviances and percent contributions for each term. The Coefficients table adds confidence intervals for the coefficients. The Fits and Diagnostics table adds the standard error of the fit, the 95% confidence interval for the fit, the deleted residual, the HI, the Cook's D, and the DFITS.

TSIMPLE
Displays the simple versions of the Analysis of Variance table, the Coefficients table, and the Fits and Diagnostics table.

TMETHOD
Displays a table of information about the method and options.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model.

If you do not specify FULL or NOFULL, then Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Specifies to display the coefficients, p-values, and model summary statistics for each step of the procedure.

NOFULL
Hides these statistics.

TFACTOR
Displays information about the factors.

TRINFO
Displays the number of events, non-events, and the total. The table also indicates which response value is the reference event.

TSTEP
Displays the deviance at each iteration of the coefficients.

TCOEFFICIENTS

Displays the coefficients and the p-values for the Wald normal approximation tests.

If you do not specify FULL or NOFULL, Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Specifies to show the coefficients for the reference level when you use (0, 1) coding for categorical predictors.

NOFULL

Hides these statistics.

TODDS

Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model.

INCREMENT K...K

Specifies the increment for the odds ratio for a covariate with K. For example, if the covariate is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram. Enter a value for each covariate in the order that the covariates follow the COVARIATE subcommand. Enter 1 for a covariate to use the same units as the data.

TSUMMARY

Displays the statistics that evaluate model fit, including deviance R^2.

TGOODNESS

Displays the Pearson and deviance goodness-of-fit tests. For the binomial distribution models, also shows the Hosmer-Lemeshow goodness-of-fit test.

TDEVIANCE

Displays the ANOVA table.

TEQUATION

Displays the regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TMEANS

Displays the table of fitted mean probabilities for all combinations of the factor levels that you specify in MEANS.

THOSMER

Displays the table of observed and expected values for the Hosmer-Lemeshow test.

TASSOCIATION

Displays the number of concordant and discordant pairs, the Somers' D, the Goodman-Kruskal Gamma, and Kendall's Tau-a.

TDIAGNOSTICS [K]

Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

Storage subcommands

FITS C

Stores the event probabilities in C.

RESIDUALS C
Stores the deviance residuals in C by default. Stores the Pearson residuals if you use the subcommand PEARSON.

SRESIDUALS C
Stores the standardized deviance residuals in C by default. Stores the standardized Pearson residuals if you use the subcommand PEARSON.

TRESIDUALS C
Stores the deleted deviance residuals in C by default. Stores the deleted Pearson residuals if you use the subcommand PEARSON.

COOK C
Stores the Cook's distances in C.

DFITS C
Stores the DFITS in C.

HI C
Stores the leverages in C.

SADJUSTEDRESP C
Stores the final adjusted response values in C.

SIWEIGHTS C
Stores the internal weights used to estimate the model parameters in C.

SECOEFFICIENTS C
Stores the standard errors of the estimated coefficients in C.

COEFFICIENTS C
Stores the estimated coefficients in C.

SMEANS C
Stores the fitted mean probability for the specified terms in C.

DBETA C
Stores the change in the estimated regression coefficients (delta b) when you delete all observations with a particular factor/covariate pattern in C. Use the delta b values to detect observations with a strong influence on the coefficients.

DSBETA C
Stores the change in the estimated regression coefficients when you delete all observations with a particular factor/covariate pattern based on the standardized Pearson residual in C.

DCHISQUARE C
Stores the change in the chi-square statistic when you delete all observations with a particular factor/covariate pattern in C. Observations that are poorly fit by the model have high delta chi-square values.

DDEVIANCE C
Stores the change in the deviance when you delete all observations with a particular factor/covariate pattern in C.

XMATRIX *M*
Stores the design matrix in M.

XPWXINVERSE *M*

Stores the X'WX inverse matrix of the estimated coefficients in M.

SPFACTORIAL: Session command for analyzing a 2-level split-plot design

SPFACTORIAL

Analyzes a 2-level split-plot designs that is produced by SPDESIGN on page 285. In a split-plot experiment, levels of the hard-to-change factor are held constant for several experimental runs, which are collectively treated as a whole plot treatment. The easy-to-change factors are varied over these runs, each of which is considered a subplot treatment.

Your split-plot design can include covariates and blocks. Specify interactions in the following way: C5*C6.

You can display effects plots (normal, half-normal, and Pareto) for either subplot effects only or subplot and whole plot effects. You can also display residual plots. You can store the residuals, fitted values, and other diagnostics for further analysis.

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

FACTOR C...C (required)

Specifies the column that contains the response variable. The column must be numeric or date/time.

HTCHANGE C...C

Specifies factors C...C as hard-to-change. These columns are a subset of the columns that follow FACTOR.

WPID C

Specifies the whole plot that each experimental run belongs in.

TERMS *termlist*

Specifies the model terms. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

BLOCKS C

Specifies the column that contains the blocking information (optional). Include the blocking column with TERMS to make a model that considers blocks.

COVARIATES C...C

Specifies columns that include covariates. Include the covariate columns with terms to make a model that considers the covariates. Covariates are fit first, the blocks second, and then all other terms. You can have up to 50 covariates.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

LEVELS K K,..., K K

LEVELS C

Specifies the uncoded levels for the factors. For each factor, the list or column must contain two numerical values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all the levels are continuous or all of the levels are categorical.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

BOXCOX [K]

Performs a Box-Cox transformation with a specified lambda. K is the value of lambda and must be between –5 and +5. If K is not given, Minitab will find the optimal lambda. By default, Minitab rounds the optimal value. To use the exact lambda, specify NOROUND.

Minitab cannot calculate the optimal lambda for stepwise regression. Consequently, you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

NOROUND

Specifies to use the exact lambda.

MEANS *termlist*

Calculates least squares means for the specified terms. Use TMEANS to display these means. Use SMEANS to store these means, or BSMEANS if you transformed the data.

TOLERANCE K

Requires that Minitab retain a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

K must be positive and less than 1. The default value is 8.88E-16.

Graphs

GEFFECTS

Displays a normal probability plot of the effects or standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GWHOLE

Specifies to show plots for whole plot effects that are in the model as well as subplot effects.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and displays the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram or an individual value plot of the residuals, depending on the sample size.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals versus fitted values, and residuals versus the order of the data, or versus the order specified in C.

GVARIABLE C...C

Displays a separate graph for the residuals versus each specified column that.

Results

The following subcommands specify what is displayed in the results.

NODEFAULT

Specifies that no default tables or graphs are displayed.

TEXPAND

Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE

Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TMETHOD

Displays the method table.

TEQUATION

The regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TCOEFFICIENTS

Displays the table of coefficients.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

The summary of model table, which includes the R^2 statistics and S. The expanded version of the table includes the prediction sum of squares (PRESS).

TANOVA

The ANOVA table. The expanded version of the table includes the sequential sums of squares and the percent contribution to the total sum of squares.

TDIAGNOSTICS [K]

A table of diagnostics. The expanded version of the table includes the standard error of the fit, the confidence interval for the fit, the deleted residual, the leverage, the Cook's D, and the DFITS.

Value of K	Diagnostics that are displayed
0	Diagnostics only for unusual observations
1	Diagnostics for all observations

TALIAS [K]

Displays alias information. One line is displayed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to three-way interaction for 7 to 10 factors, and two-way interactions for more than 10.

Note A value of K that is larger than the default can take a long time to calculate.

TMEANS

The least squares means for all combinations of the factor levels that you specify in MEANS.

Storage

The following subcommands specify what is stored in the worksheet.

FITS C

Stores the fitted values.

BFITS

Stores the back-transformed fits after the Box-Cox transformation.

BCRESP C

Stores the Box-Cox transformation of the response.

RESIDUALS C

Stores the residuals (response – fit)

SRESIDUALS C

Stores the standardized residuals.

TRESIDUALS C

Stores the deleted Studentized residuals.

WRESIDUALS C...C

Stores thole plot residuals.

EFFECTS C

Stores the effects that are included in the output.

Note Effects are not printed or stored for the constant, covariates, or blocks.

COEFFICIENTS C

Stores the coefficients that are included in the coefficients table.

The coefficients are based on coded levels (–1, 0, +1), not on the actual levels in the data set. A coefficient is always half the value of the corresponding effect.

XMATRIX *M*

Stores the design matrix. The first column is a column of ones for the constant term. Next, there is one column for each covariate. Next, if the design was blocked into k blocks, there are (k – 1) columns. Next, there is one column for each factor, using –1 for the low level, 0 for a center point, and +1 for the high level. Next, there is one column for each interaction in the model. The column for an interaction is the product of all the columns that correspond to the factors that are contained in that interaction.

If SPFACTORIAL removed some of the interactions that you specified in your model, those interactions do not appear in the stored design matrix. The columns in XMATRIX match the coefficients that are displayed and stored.

WXMATRIX *M*

WXMATRIX is like XMATRIX, but it includes the whole plot error terms.

WCOEFFICIENTS C...C

Stores coefficients in C...C (including whole plot error terms).

HI C

Stores the leverage values that you use to identify influential observations.

SMEANS C

Stores the least squares means from MEANS in the specified column.

If you use a Box-Cox transformation, SMEANS stores the means for the transformed response.

BSMEANS C

If you use a Box-Cox transformation, BSMEANS stores the means for the back-transformed response.

VPREPROCESS: Session command to preprocess responses for analyze variability

Note VPREPROCESS replaces the obsolete session command PREPROCESS.

VPREPROCESS

Use VPREPROCESS to calculate and store standard deviations and counts of repeat or replicate measurements before you use VFACTORIAL on page 328 to analyze the variability of responses. Use one of the following subcommands: REPEATS, REPLICATES, or DEFINE. Then use the sub-subcommands to specify how to preprocess the data.

REPEATS

Specifies that repeats of the measurement are in the same row of different columns.

RESPONSE C...C

Specifies which columns contain the repeated measurements of the response variable.

STDEV C

Specifies where to store the standard deviations of the repeated measurements for analysis.

COUNT C

Specifies where to store the number of repeated measurements for analysis.

RMEAN C

Specifies where to store the mean of the repeated measurements for analysis.

REPLICATES

Specifies that replicated measurements of a response variable are in the same column.

RESPONSE C...C

Specifies which columns contain the replicated measurements of a response variable. Each column is a separate response variable.

STDEV C...C
Specifies where to store the standard deviations of the replicate measurements for analysis. Enter a column for each response variable.

COUNT C...C
Specifies where to store the number of replicated measurements for analysis. Enter a column for each response variable.

COVARIATES C...C
Specifies covariates to adjust for in the calculation of the standard deviations.

DEFINE C...C
Specifies that standard deviations for analysis are already in the worksheet. Each column contains the standard deviations for a different response variable.

COUNT E...E
Specifies the number of measurements for the standard deviation in each row. If the numbers of measurements are all the same, enter the number or enter a column of constant values. If the numbers of measurements vary, use a column to indicate the number for each row.

VFACTORIAL: Session command for analyzing variability in a 2-level factorial design

VFACTORIAL
Analyzes variability in a 2-level factorial design, including those produced by FFDESIGN on page 280 and PBDESIGN on page 259. You can specify covariates and a blocking factor. You can fit one response at a time. Before you use VFACTORIAL, you must calculate and store standard deviations using PREPROCESS on page 327. Minitab uses least squares estimation by default, but you can specify maximum likelihood estimation using MLE.

The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, Minitab defines the low level as the level that is first using alphabetical order. Change the order of the levels after the LEVELS subcommand to change this default ordering.

Optionally, you can display seven plots: a normal plot of the effects, a half normal plot of the effects, a Pareto chart of the effects, and four different residual plots. In addition, you can store the residuals, fitted values, and many other diagnostics for further analysis.

RESPONSE C
Specifies the column that contains the standard deviations.

COUNT E (required)
Specifies the constant or column that contains the number of repeat or replicate measurements. The columns must have the same number of rows that the response column has.

FACTOR C (required)
Specifies the factors. The columns must have the same number of rows that the response column has.

BLOCKS C
Specifies the column that contains the blocking information (optional). A blocking variable column can be numeric or text, and can contain any value. The levels do not need to be in any special order. When the

blocking column is numeric, the smallest value is block 1, the second smallest value is block 2, and so on. Include the blocking column with TERMS to make a model that considers blocks.

CTPT C
Specifies the column that identifies which design points are center points. Include the center point column with TERMS to make a model that tests for curvature.

TERMS C
Specifies the model terms. Only continuous predictors can repeat in a single term. Nested terms are not allowed. The model can be nonhierarchical.

INUNIT K
Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

LEVELS
Specifies the uncoded levels for the factors. For each factor, the list or column must contain two numerical values. Use LEVELS to change the default order of levels that Minitab uses. You can use a column only when all the levels are continuous or when all the levels are categorical.

MLE
Specifies maximum likelihood estimation. By default, Minitab uses least squares estimation to fit the model.

Covariates

COVARIATES C...C
Lists the covariates to include in the model. You can have up to 50 covariates.

Options

MLE
Specifies using maximum likelihood estimation to analyze the model. The default method is least squares estimation.

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K
Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

MEANS *termlist*

Calculates least squares means for the specified terms. You can use the following commands to display and store the least squares means. If the design is orthogonal and there are no covariates, each mean is the average of all the observations in the corresponding cell. Use the same syntax and abbreviations on MEANS as you use in specifying a model.

TMEANS

Displays the calculated least squares means for the specified terms.

SMEANS

Stores the calculated least squares means for the specified terms.

BSMEANS

If you transformed the data, stores the calculated least squares means for the specified terms.

TOLERANCE K

Requires that Minitab retain a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

K must be positive and less than 1. The default value is 8.88E-16.

Stepwise

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

Once Minitab fits the full model or a model that leaves 1 degree of freedom for error, the procedure stops. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

GEFFECTS

Displays a normal probability plot of the effects or standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER [C]
Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]
Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C
Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

The following subcommands specify what is displayed in the results.

NODEFAULT
Specifies that no default tables or graphs are displayed.

TEXPAND
Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE
Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TMETHOD
Displays the method table.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Displays the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

NOFULL
Does not display the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

TEQUATION
The regression equation table. If the model is hierarchical, the equation is in uncoded units. If the model is nonhierarchical, the equation is in coded units.

TCOEFFICIENT

Displays the table of coefficients. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

The summary of model table, which includes the R^2 statistics and S. The expanded version of the table includes the prediction sum of squares (PRESS), the corrected Akaike's Information Criterion (AICc) and the Bayesian Information Criterion (BIC).

TANOVA

The ANOVA table. The expanded version of the table includes the sequential sums of squares and the percent contribution to the total sum of squares.

TDIAGNOSTICS [K]

Displays a table of diagnostics.

Value of K	Diagnostics that are displayed
0	Diagnostics only for unusual observations
1	Diagnostics for all observations

If you do not specify K, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

TALIAS [K]

Alias information. One line is printed for each term in the model, with a list of all terms, up to order K, that are aliased with the model term. If K is omitted, Minitab uses the default rule of FFDESIGN: all interactions for 2 to 6 factors, up to 3-way interaction for 7 to 10 factors, and 2-way interactions for more than 10.

TMEANS

The least squares means for all combinations of the factor levels that you specify in MEANS.

Storage

The following subcommands specify what is stored in the worksheet.

LFITS C

Stores the fitted values for the natural log of the standard deviation.

RRESIDUALS C

Stores the ratio residual, which is the ratio of the response standard deviation to the fitted standard deviation.

LRESIDUALS C

Stores the log residual, which is the difference between the natural log of the response standard deviation and the natural log of the fitted standard deviation.

SLRESIDUALS C

Stores the standardized log residual, which is the log residual divided by its (asymptotic) standard error.

EFFECTS C
Stores the effects that are included in the output.

Note Effects are not printed or stored for the constant, the covariates, or the blocks.

REFFECTS C
Stores the ratio effects.

COEFFICIENTS C
Stores the coefficients that are included in the coefficients table.

The coefficients are based on coded levels (–1, 0, +1), not on the actual levels in the data set. A coefficient is always half the value of the corresponding effect.

XMATRIX *M*
Stores the design matrix.

HI C
Stores the leverage values.

COOK C
Stores Cook's distance.

DFITS C
Stores the DFITS (also called DFFITS) values.

LNRESP C
Stores the log of the standard deviations. This column is the response for the model.

BFITS
Stores the back-transformed fitted values for the standard deviations.

SMEANS C
Stores the means for terms that are specified in the MEANS subcommand in the specified columns.

BSMEANS C
If you use a Box-Cox transformation, BSMEANS stores the back-transformed means.

SWEIGHTS C
Stores the unadjusted weights based on the fitted variances to use when you analyze location effects with FFACTORIAL.

SADJUSTED C
If your data contain repeats with some replicated points, SADJUSTED stores the adjusted weights.

RMEANS C
Specifies the row means that are associated with the row standard deviations used as the response.

ACOVARIATES C...C
Accounts for covariates in C...C in calculating adjusted weights.

Plot Factorial Design

FACPLOT: Session command for creating a factorial plot

FACPLOT C

FACPLOT "T|K"

Displays factorial plots for the model associated with the response variable in the specified column. To create this graph, you need to specify a model. To produce an interaction plot, you must have two or more variables.

Displays factorial plots for the model that you specify. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following code: `FACPLOT 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `FACPLOT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

To produce an interaction plot, you must have two or more variables. These plots can illustrate how a response variable relates to one or more variables. You can produce two types of factorial plots.

- Use a main effects plot to compare fitted means or data means across levels of one or more variables.
- Use an interactions plot to assess the effects of one variable at levels of another variable. An interactions plot displays the fitted means for a level of a variable with the level of a second variable held constant.

Although you can use these plots to display the effects, be sure to evaluate significance by looking at the effects in the analysis of variance table.

FACTORS C...C

Specifies the variables to display in both the main effect and interaction plots. You must specify 2 or more variables to generate an interaction plot.

GMAIN

GMAIN displays the main effect plot.

GINT

GINT displays the interactions plot.

FULL

Use FULL to display both interaction plots for each pair of categorical variables, such as Predictor 1 by Predictor 2 and Predictor 2 by Predictor 1. Sometimes you can identify interactions more easily when you examine both plots.

LOWER

Use LOWER to display one version of each interaction plot.

GMODEL (default)

GMODEL display plots for only terms that are in the model and listed in FACTORS. For purposes of a main effect plot, a variable is in the model if the model includes any terms that involve only that variable, e.g. A, A*A, or A*A*A. For purposes of an interaction plot, the interaction of two variables is in the model if the model includes any terms that involve both variables, e.g. A*B, A*B*B.

GALL

GALL displays plots for all terms involving variables on the FACTORS subcommand.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MFFCUBE: Session command for creating a cube plot for fitted means

MFFCUBE C

Creates a cube plot of the fitted means for the response in C.

You can draw a cube plot for a model from a 2-level factorial design. The factor columns can be numeric or text, and can contain any values. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2. All factors must have 2 levels, and can have center points.

To draw a cube plot of the data means or a cube plot of the process variables in a mixtures design, use FFCUBE on page 341.

FACTORS C...C

Specifies the factors to include on the cube plot in C...C. You can specify 2–8 factors.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

FFINT: Superseded by FACPLOT except for mixture designs

Note FFINT is superseded by FACPLOT on page 277, except for mixture designs.

FFINT C...C

Creates interaction plots for factors in C...C.

Creates an interaction plot for 2-level factorial designs. If you enter 2 factors on the command line, then Minitab displays a single interaction plot. If you enter more than 2 factors, then Minitab displays a matrix of interaction plots. You can use factors from designs that are generated by FFDESIGN and PBDESIGN.

You can specify up to 15 columns that contain the factors. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on. All factors must have 2 levels, and can have center points.

RESPONSE C (required)

Plots the response data in C. If you specify RESPONSE without FMEANS, then Minitab plots the raw response data. The interaction effects are calculated based on the means of the data. If you plot process variables from a mixture design, then Minitab fits only the raw response data, not the fitted means.

You can specify up to 15 factors.

FMEANS C...C

Plots the least squares or fitted means in C...C. If you specify RESPONSE with FMEANS, then Minitab plots the predicted values from the model. The interaction effects are calculated based on the fitted means. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

FULL

Displays the full interaction matrix when you enter more than 2 factors. By default, Minitab displays the half matrix. The half matrix displays the plots in the upper right portion of the matrix. In the full matrix, the transpose of each plot in the upper right is displayed in the lower left portion of the matrix. The full matrix takes longer to display than the half matrix.

YMIN K

Specifies the minimum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

YMAX K

Specifies the maximum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

CTPT C

Specifies the center point column in C. Minitab considers a value of 0 as a center point.

LEVELS K K, ..., K K

Specifies the levels for the uncoded units for each factor in K K, ..., K K. Use LEVELS to change the default order of levels that Minitab uses.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, which indicates that the factors are process variables.

FFMAIN: Superseded by FACPLOT except for mixture designs

Note FFMAIN is superseded by FACPLOT on page 277, except for mixture designs.

FFMAIN C...C
Creates a main effects plot for 2-level factorial designs, for factors in C...C.

FFMAIN plots the means at the two levels of each factor and connects them with a line. FFMAIN plots any center points you have, using a different symbol than that used for the factorial points. A reference line is drawn through the plot to represent the overall mean (also called grand mean) of the response data.

Specify up to 50 factors. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1. The next entry that is different is level 2. All factors must have 2 levels, and can have center points.

RESPONSE C (required)

Plots the response data in C. If you specify RESPONSE without FMEANS, then Minitab plots the raw response data. The main effects are calculated based on the means of the data. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

With RESPONSE, you can specify up to 50 factors.

FMEANS C...C

Plots the least squares or fitted means in C...C. If you specify RESPONSE with FMEANS, then Minitab plots the predicted values from the model. The main effects are calculated based on the fitted means. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

With FMEANS, you can specify up to 10 factors.

YMIN K

Specifies the minimum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

YMAX K

Specifies the maximum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

CTPT K

Specifies the center point column in K. Minitab considers a value of 0 as a center point.

LEVELS K K, ..., K K

Specifies the levels for the uncoded units for each factor in K K, ..., K K. Use LEVELS to change the default order of levels that Minitab uses.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, which indicates that the factors are process variables.

FFCUBE: Session command for creating a cube plot

FFCUBE C...C

Use FFCUBE to draw a cube plot that does not show mean response values. With a mixture design, you can use FFCUBE to draw a cube plot that shows mean response values. For all other models, such as the models from FFAC or REGR, use MFFCUBE on page 337 to draw a cube plot that shows mean response values.

Displays a cube plot for factors in C...C. You can draw a cube plot with 2–8 factors for each response entered with the RESPONSE subcommand with data means or 2–7 factors with fitted means. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on. All factors must have 2 levels, and can have center points.

When you use the RESPONSE subcommand, FFCUBE displays the means for the response at each point in the cube where observations are measured. When you do not use the RESPONSE subcommand, FFCUBE draws points on the cube for the effects that are in your model, which helps you visualize the design.

RESPONSE C...C
Plots the raw response data that is in columns C...C. The means for the cube points are calculated from the data. You can specify 2–8 factors.

If you use FMEANS, RESPONSE is required.

FMEANS C...C

Plots the least squares or the fitted means that are in columns C...C. You can specify 2-7 factors. The column must contain the fitted means in standard order for the cube points. The last row must be the overall mean of the data.

If you use FMEANS, RESPONSE is required.

INUNIT K (required)

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded form
1	Uncoded form

CTPT C

Specifies a center point column in C. Minitab considers a value of 0 to be a center point. A valid column contains only values of 0 and 1.

If you don't use MIXTURE, CTPT is required. You cannot use both MIXTURE and CTPT.

LEVELS K K, ..., K K (required)

Specifies the levels for the uncoded units for each factor in K K, ..., K K. The order of the parameters must be identical to the order of the factors that follow FFCUBE.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, indicating that the factors are process variables.

Response Surface Designs

BBDESIGN: Session command for creating a Box-Behnken design

BBDESIGN K [K]
Generates Box-Behnken designs for 3, 4, 5, 6, 7, 9, or 10 continuous factors and optional categorical factors.

Enter the number of factors on BBDESIGN. For a summary of the designs generated by BBDESIGN, go to Box-Behnken designs on page 1218. All the blocked designs have orthogonal blocks.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.)

CENTER K
Adds the number of center points to the design that you specify in K, up to 50. If you do not use this subcommand, then Minitab uses a default number of center points, based on the design you specified.

If the design is blocked, the center points are divided equally (as much as possible) among the blocks.

REPS K
Replicates each point the number of times that you specify in K, up to 50 replicates of each point.

BLOCKS K
Specifies the number of blocks in K. The choices for K are limited. The "Optional number of blocks" column in Box-Behnken designs on page 1218 shows which values of K you can use for different designs.

RBLOCK

Use to specify blocks on replicates. Each set of replicate points is placed in a separate block. No arguments are taken.

LEVELS K K, ..., K K

LEVELS C

LEVELS allows you to override the default values normally assigned to the levels of each factor. The default values use -1 for the low level of the corner point, +1 for the high level, and 0 for the center point.

With LEVELS K K, ..., K K, you list a pair of arguments for each factor. The first K should be the low level and the second K should be the high level. Levels must be numeric.

With LEVELS C, put the level values in a column, in the same order as if they were listed on the subcommand. Then specify the level column on the LEVELS subcommand.

CNUM K...K

Specifies the number of levels in categorical factors. The number of arguments must match the second argument for the main command. Each K must be an integer from 2 to 100. The default number of levels in categorical factors is 2.

CLEVEL K...K

CLEVEL C

Specifies the uncoded values of the categorical factor levels in K...K or stored in C. The number of arguments K must equal the sum of arguments in CNUM. The group of arguments for each factor must be of the same type. If stored in column C, the column can be numerical or text and must equal the sum of arguments in CNUM. The default categorical factor levels are "1", "2", "3", and so on.

RANDOMIZE [K]

Randomizes the run order and optionally uses base K. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output for the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.

Value of K	Output that is displayed
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

CCDESIGN: Session command for creating a central composite design

CCDESIGN K

Creates a central composite design with K number of continuous factors and optional K categorical factors.

Generates central composite designs for 2 to 10 continuous factors and optional categorical factors.

Specify the number of factors on CCDESIGN. Designs for 1 to 7 factors can use either a full or fractional factorial design. Designs with 5 to 8 factors can use a one-half fraction of the full factorial portion of the central composite design. Use HALF to specify the one-half fraction. Designs with 8 or 9 factors can use a one-quarter fraction of the full factorial portion of the central composite design. Use QUARTER to specify one-quarter fraction. Designs with 10 factors can use a one-eighth quarter fraction of the full factorial portion of the central composite design. Use EIGHTH to specify one-eighth fraction.

Central composite designs on page 1219 summarizes the designs generated by CCDESIGN. The default values of alpha provide orthogonal blocking and, whenever possible, rotatability. To generate a face-centered design, use ALPHA = 1. The cube portions of these designs are identical to those generated by FFDESIGN on page 280 with the same number of center points and blocks. Thus, a design generated by FFDESIGN with the same number of runs, center points, and blocks as shown below can be built up into an orthogonally blocked central composite design. Any FFDESIGN with the right number of runs and blocks can be built up into a blocked CCDESIGN. However, to make the blocks orthogonal the FFDESIGN must use the number of center points shown in the table.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.).

HALF

Specifies the one-half fraction of a full factorial design. All the default designs generated by CCDESIGN use a full factorial in the cube portion of the design. With five or six factors, you can optionally use a one-half fraction of the full factorial. Use HALF to do so.

QUARTER

Specifies the one-quarter fraction of a full factorial design. Use QUARTER if you have eight or nine factors and want to use a one-quarter fraction of the full factorial.

EIGHTH

Specifies the one-eighth fraction of a full factorial design. Use EIGHTH if you have ten factors and want to use a one-eighth fraction of the full factorial.

CENTER K [K]

Adds K center points to cube and optionally, K center points to the axial block. If you do not give this subcommand, Minitab uses a default number of center points, based on the design you specify.

If the design is blocked, you must give both Ks. The first gives the number of center points in the cube and the second gives the number of center points in the axial block. If the cube is blocked, the number of center points in the cube is split equally (as much as possible) among the blocks.

Central composite designs on page 1219 shows the defaults for blocked and unblocked designs. The defaults for unblocked designs are in the "Total center points" column. For blocked designs, the defaults are in the "Cube center points" and "Axial center points" columns.

ALPHA K

Specifies an alpha value in K. CCDESIGN has a default value for alpha, shown in Central composite designs on page 1219, for each design. This subcommand allows you to change the default. The value must be positive. A value of 1 results in a face-centered design. A value that is less than 1 places the axial points inside the cube. A value that is greater than 1 places the axial points outside the cube.

REPS K

Replicates each point the number of times that you specify in K, up to 50 replicates of each point.

BLOCKS K

Specifies number of blocks in K. The choices for K are limited. The "Total blocks" column in Central composite designs on page 1219 shows which values of K you can use for different designs.

RBLOCK

Blocks on replicates. Each set of replicate points is placed in a separate block. No arguments are taken.

LEVELS K K, ..., K K

LEVELS C

LEVELS allows you to override the default values normally assigned to the levels of each factor. The default values use -1 for the low level of the corner point, +1 for the high level, 0 for the center point, and plus and minus alpha for the axial points.

With LEVELS K K, ..., K K, you list one pair of arguments—giving first the low level and then the high level for each factor. The center point and star points are scaled accordingly. Levels must be numeric.

With LEVELS C, put the level values in a column, in the same order as if they were listed on the subcommand. Then specify this level column on the LEVELS subcommand.

One important thing to note is that the values you enter in LEVELS are usually not the minimum and maximum values in the design. They are the low and high settings for the cube portion of the design. The axial points are usually outside the cube (unless you specify an alpha that is less than or equal to 1). If you are not careful, this could lead to axial points that are not in the region of interest, or that may be impossible to run.

If you wish, you can use LEVELS to guarantee that all of the design points will fall within a predetermined minimum and maximum value for the factors. What you need to determine are the appropriate low and high settings for the cube, so you can enter these into LEVELS. All you need to do so are the desired minimum and maximum and the value of alpha for the design. Default alpha values are shown in Central composite designs on page 1219.

CNUM K...K

Specifies the number of levels in categorical factors. The number of arguments must match the second argument for the main command. Each K must be an integer from 2 to 100. The default number of levels in categorical factors is 2.

CLEVELS K...K

CLEVELS C

Specifies the uncoded values of the categorical factor levels in K...K or stored in C. The number of arguments K must equal the sum of arguments in CNUM. The group of arguments for each factor must be of the same type. If stored in column C, the column can be numerical or text and must equal the sum of arguments in CNUM. The default categorical factor levels are "1", "2", "3", and so on.

INSCRIBE

Defines the low and high pairs in LEVELS as the axial points instead of the cube points.

RANDOMIZE [K]

Randomizes the run order and optionally uses base K. If you specify blocks, randomization is done separately within each block and then the blocks are randomized. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random number generator and can take any integer as an argument. To get the same ordering of the rows, create the same design and randomize the runs with the same base. For a discussion, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output from the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

RSREG: Session command for analyzing a response surface design with least squares regression

RSREG

Analyzes a response surface design that can have continuous or categorical predictors using least squares regression.

Fits a response surface model. Designs generated with BBDESIGN on page 343 and CCDESIGN on page 345 can be analyzed using this command. Specify interactions in the following way: C5*C6. You can only use extra text on the command and subcommand lines after a #. For more information on specifying the model, go to How to specify the model for response surface designs on page 1251.

You can have up to 9 factors. To block the design, specify a blocking variable with the BLOCKS subcommand. Optionally, you can display any of the five residual plots, and store the residuals, fitted values, and diagnostics for further analysis.

XMATRIX is useful for storing the design matrix for use in other Minitab commands. It takes the factors, creates the squares and cross-products, and stores all of these in a matrix. Copy the matrix into columns for use in other commands.

You can store the residuals, fitted values, and many other diagnostics for further analysis.

For information on how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

RESPONSE C

Specifies the column that contains the response variable. The column must be numeric or date/time.

BLOCK C

Specifies the column that contains the blocking information (optional). If you use BLOCK, you must also include the blocking column in TERMS.

FACTOR C...C

Is mandatory and specifies the continuous factors. The columns must be numeric or date/time and must match the length of the response column.

CFACTOR C...C

Specifies the categorical factors if you have any. The columns can be numeric, text, or date/time and must match the length of the response column.

CNUM K...K

Use both CNUM and CLEVELS when the categorical factor levels in the worksheet are coded and you want to display uncoded levels in the results.

CNUM specifies the number of levels for each categorical factor. Each K must be an integer from 2 to 100. You must specify the number of K arguments as the number of columns that you specify in CFACTOR.

CLEVELS *E...E*

Use both CNUM and CLEVELS when the categorical factor levels in the worksheet are coded and you want to display uncoded levels in the results.

CLEVELS specifies the uncoded values for all the categorical factor levels. You must specify the same number of arguments as the number of K arguments that you specify in CNUM. The group of arguments for each factor must be of the same type.

LEVELS C

LEVELS K K, ..., K K

Specifies uncoded levels for all continuous factors in C or K...K. For each continuous factor, the list or column must contain two numerical values. Use LEVELS to change the default order of levels that Minitab uses.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

TERMS *termlist*

Specifies the terms to include in the model. Terms can contain a maximum of two factors. Only continuous predictors may be repeated. Nested terms are not allowed. Terms must contain the blocks if they are to be fitted. The model can be nonhierarchical.

Options

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$$(\mathbf{X}'\ \mathbf{W}\ \mathbf{X})^{-1}\ (\mathbf{X}'\ \mathbf{W}\ \mathbf{Y})$$

This is equivalent to minimizing the weighted sum of squares for error:



The prediction interval computed by PREDICT assumes a weight of 1. If you used the WEIGHT subcommand with values other than 1, you should adjust the prediction interval values manually.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

TOLER K

Specifies tolerance level for collinearity and constant check. By default, K = 2.2204460492503131e–016.

Box-Cox transformation

BOXCOX [K]

Performs a Box-Cox transformation with a specified lambda. K is the value of lambda and must be between –5 and + 5. If K is not given, Minitab will find the optimal lambda. By default, Minitab rounds the optimal value.

Minitab cannot calculate the optimal lambda for stepwise regression. Consequently, you must specify a lambda value for BOXCOX if you use STEPWISE, FORWARD, or BACKWARD.

Stepwise regression

`FINFORMATION`, `STEPWISE`, `FORWARD`, and `BACKWARD` perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is 1 – the confidence level that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of errors. This plot lets you look at both the magnitude and the importance of an effect at the same time. The alpha level is 1– the confidence level that follows CONFIDENCE, unless you use stepwise selection. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K

Specifies to include the model terms and the terms that are not in the model up to order K on the plots. K can be 1 or 2. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL

Specifies that only the terms that are in the model are on the plots.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals versus fitted values, and residuals versus order of the data.

GVARIABLE C...C

Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

NODEFAULT

Specifies that no default tables or graphs are displayed.

TMETHOD

Displays the method table.

TMSDETAILS

Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL

Displays the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

NOFULL

Does not display the coefficients, the p-values, Mallows' Cp, and the model summary statistics for each step of the procedure.

TEQUATION

Displays the regression equation table. Minitab displays up to 50 equations. If you do not specify SINGLE or SEPARATE, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

SINGLE

Displays a single equation, rather than a separate equation for each factor level combination.

SEPARATE

Displays separate equations.

TCOEFFICIENT

Displays the table of coefficients. If you do not specify FULL or NOFULL, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL

Displays the full set of coefficients for categorical predictors.

NOFULL

Displays only the linearly independent coefficients.

TSUMMARY

Displays the model summary table.

TANOVA

Displays the ANOVA table.

TDIAGNOSTICS [K]

Displays a table of diagnostics.

Value of K	Diagnostics that are displayed
0	Diagnostics only for unusual observations
1	Diagnostics for all observations

If you do not specify K, then Minitab uses the settings in **File** > **Options** > **Linear Models** > **Display of Results**.

TEXPAND

Displays the expanded versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

TSIMPLE
Displays the simple versions of the ANOVA table, the table of coefficients, the model summary table, and the table of unusual observations.

Storage

BCRESP C
Stores the Box-Cox transformation of the response.

BFITS C
Stores the fits for the original response.

FITS C
Stores the fitted values.

RESIDUALS C
Stores the residuals (fitted values – observed values).

SRESIDUALS C
Stores the standardized residuals.

TRESIDUALS C
Stores the deleted Studentized residuals in C.

HI C
Stores the leverage values.

COOK C
Stores Cook's distance.

DFITS C
Stores the DFITS (also called DFFITS) values.

COEFFICIENTS C
Stores the estimated coefficients.

XMATRIX *M*
Stores the design matrix for the regression model in M.

BRSREG: Session command for analyzing a response surface design with a binary response

Required commands

To perform the analysis, use the main command and at least these accompanying subcommands.

BRSREG
Analyzes a response surface design that can have continuous or categorical predictors with a binary response.

Fits a response surface model. Designs generated with BBDESIGN on page 343 and CCDESIGN on page 345 can be analyzed using this command. Specify interactions in the following way: C5*C6. For more information on specifying the model, go to How to specify the model for response surface designs on page 1251.

You can have up to 9 factors. To block the design, specify a blocking variable with the BLOCKS subcommand. Optionally, you can display any of the five residual plots, and store the residuals, fitted values, and diagnostics for further analysis.

XMATRIX is useful for storing the design matrix for use in other Minitab commands. It takes the factors, creates the squares and cross-products, and stores all of these in a matrix. Copy the matrix into columns for use in other commands.

You can store the residuals, fitted values, and many other diagnostics for further analysis.

For information on how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

RESPONSE C C

Enters the response data as a column that contains the number of events first and then a column that contains the number of trials. The columns must have the same number of rows. For examples, go to Entering data for response variables on page 1225.

FACTOR C...C

Specifies the continuous factors. The columns must be numeric or date/time and must have the same number of rows as the response columns. Specify any categorical factors after CFACTOR instead of FACTOR.

TERMS *termlist*

Specifies the terms to include in the model. Terms can contain a maximum of two factors. Only continuous predictors may be repeated. Nested terms are not allowed. Terms must contain the blocks if they are to be fitted. The model can be nonhierarchical.

LEVELS C or LEVELS K K, ..., K K

Specifies uncoded levels for all continuous factors in C or K...K. For each continuous factor, the list or column must contain two numerical values. Use LEVELS to change the default order of levels that Minitab uses. Specify levels for categorical factors after CLEVELS instead of LEVELS. Also specify the number of levels for categorical factors after CNUM.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

Optional subcommands

Use these subcommands and the subcommands in the following sections to change the default analysis.

CFACTOR C...C

Specifies the categorical factors if you have any. The columns can be numeric, text, or date/time and must match the length of the response column.

CNUM K...K

Use both CNUM and CLEVELS when the categorical factor levels in the worksheet are coded and you want to display uncoded levels in the results.

CNUM specifies the number of levels for each categorical factor. Each K must be an integer from 2 to 100. You must specify the number of K arguments as the number of columns that you specify in CFACTOR.

CLEVELS *E...E*

Use both CNUM and CLEVELS when the categorical factor levels in the worksheet are coded and you want to display uncoded levels in the results.

CLEVELS specifies the uncoded values for all the categorical factor levels. You must specify the same number of arguments as the number of K arguments that you specify in CNUM. The group of arguments for each factor must be of the same type.

BLOCKS C

Specifies the column that contains the blocking information in C. Include the blocking column together with TERMS to make a model that considers blocks.

EVNAME K

Specify the name of the event with K. Enclose text values in double quotation marks.

LOGIT

Use the logit link function.

NORMIT

Use the normit link function.

GOMPIT

Use the gompit link function. The gompit link function is also called the complementary log-log link function.

Options subcommands

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SEQDEVIANCE

Use sequential deviances for the tests of significance in the ANOVA table. The order of the terms in TERMS affects the statistical significance. The default is to use adjusted deviances.

PEARSON

Specifies the use of Pearson residuals. The default is to use deviance residuals.

WEIGHT C

Performs a weighted regression. An n x n matrix **W** is formed with the column of weights as its diagonal and zeros elsewhere. The regression coefficients are estimated by:

$(\mathbf{X}' \mathbf{W} \mathbf{X})^{-1} (\mathbf{X}' \mathbf{W} \mathbf{Y})$

HGROUP K

Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test with K. The default is 10.

ITERATION K

Specifies the maximum number of iterations for the optimization algorithm that estimates the coefficients. The default is 50.

TOLERANCE K

Specifies the value in K above which Minitab removes highly correlated predictors and nearly constant predictors. By default, K = 1E-12.

The tolerance value determines when Minitab retains a factor in the model that is either highly correlated with a different factor or that is almost constant. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model.

Collinearity	$1 - R^2$, where R^2 is the value that results from regressing one predictor on the remaining predictors. If $(1 - R^2)$ is < (K), then the predictor is removed from the equation.
Almost constant	If the coefficient of variation of a predictor (the standard deviation / the mean) is < square root (K), then the predictor is considered essentially constant and is removed from the equation.

CTOLERANCE K

Specifies the convergence criterion for the algorithm that estimates the coefficients. By default, K = 1E-8.

START C

Specifies a column that contains starting values for the coefficients in the optimization algorithm. The column begins with the estimate of the constant, then follows the order of the model after TERMS.

Stepwise regression subcommands

`FINFORMATION`, `STEPWISE`, `FORWARD`, and `BACKWARD` perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find a new minimum of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS

Specifies that the model is hierarchical at every step.

SINGLE

Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE

Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Graphs subcommands

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GEFFECTS

Displays a normal probability plot of the effects or the standardized effects. Minitab labels the effects on the probability plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GHALF

Displays a half normal probability plot of the effects or standardized effects. Minitab labels the effects on the half normal plot that are considered significant. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GPARETO

Displays a Pareto chart of the absolute effects. GPARETO draws a vertical reference line on the plot at the margin of error. This plot lets you to look at both the magnitude and the importance of an effect at the same time. The alpha level is (1 – the confidence level) that follows CONFIDENCE, unless you use stepwise selection that depends

on an alpha level. For forward selection, the alpha level is the level that follows AENTER. For backward and stepwise selection, the alpha level is the level that follows AREMOVE.

GALLEFFECTS K
Specifies to include the model terms and the terms that are not in the model up to order K on the plots. If you specify a stepwise method, Minitab ignores K and shows the terms in the candidate list.

GMODEL
Specifies that only the terms that are in the model are on the plots.

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Displays residuals versus fitted values.

GORDER [C]
Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]
Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C
Displays the residuals versus the variables in C. Typically, you enter the columns for the predictors.

Results subcommands

NODEFAULT
Hides default tables and graphs that you do not include the subcommand to display.

TEXPAND
Displays the expanded versions of the tables that have them. The expanded Analysis of Variance table adds the sequential deviances and percent contributions for each term. The Coefficients table adds confidence intervals for the coefficients. The Fits and Diagnostics table adds the standard error of the fit, the 95% confidence interval for the fit, the deleted residual, the HI, the Cook's D, and the DFITS.

TSIMPLE
Displays the simple versions of the Analysis of Variance table, the Coefficients table, and the Fits and Diagnostics table.

TMETHOD
Displays a table of information about the method and options.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model.

If you do not specify FULL or NOFULL, then Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Stepwise**.

FULL
Specifies to display the coefficients, p-values, and model summary statistics for each step of the procedure.

NOFULL
Hides these statistics.

TRINFO
Displays the number of events, non-events, and the total. The table also indicates which response value is the reference event.

TSTEP
Displays the deviance at each iteration of the coefficients.

TCOEFFICIENTS
Displays the coefficients and the p-values for the Wald normal approximation tests.

If you do not specify FULL or NOFULL, Minitab uses the preferences that you set in **File** > **Options** > **Linear Models** > **Display of Results**.

FULL
Specifies to show the coefficients for the reference level when you use (0, 1) coding for categorical predictors.

NOFULL
Hides these statistics.

TODDS
Displays odds ratios for each continuous predictor and each level of the categorical predictors. The output does not show odds ratios for predictors with interactions in the model.

INCREMENT K...K
Specifies the increment for the odds ratio for the predictor with K. For example, if the predictor is mass in grams, enter 1000 to see the change in the odds ratio for a kilogram. Enter a value for each numeric factor in the same order that the factors follow the FACTOR subcommand. Enter 1 for a factor to use the same units as the data.

TSUMMARY
Displays the statistics that evaluate model fit, including deviance R^2.

TGOODNESS
Displays the Pearson and deviance goodness-of-fit tests. For the binomial distribution models, also shows the Hosmer-Lemeshow goodness-of-fit test.

TDEVIANCE
Displays the ANOVA table.

TEQUATION
Displays the model equation.

SINGLE
Specifies to show one equation instead of a separate equation for each combination of the levels of the categorical predictors. The equation has a coefficient for each level of the categorical variables.

SEPARATE
Specifies to show separate equations for each combination of the levels of the categorical predictors. Each equation has a different constant term.

THOSMER
Displays the table of observed and expected values for the Hosmer-Lemeshow test.

TASSOCIATION
Displays the number of concordant and discordant pairs, the Somers' D, the Goodman-Kruskal Gamma, and Kendall's Tau-a.

TDIAGNOSTICS [K]
Displays a table of diagnostics. K = 0 displays diagnostics for only observations with high leverage values or standardized residuals greater than 2. K = 1 displays diagnostics for all observations.

Storage subcommands

FITS C
Stores the event probabilities in C.

RESIDUALS C
Stores the deviance residuals by default in C. Stores the Pearson residuals if you use the subcommand PEARSON.

SRESIDUALS C
Stores the standardized deviance residuals by default in C. Stores the standardized Pearson residuals if you use the subcommand PEARSON.

TRESIDUALS C
Stores the deleted deviance residuals by default in C. Stores the deleted Pearson residuals if you use the subcommand PEARSON.

COOK C
Stores the Cook's distances in C.

DFITS C
Stores the DFITS in C.

HI C
Stores the leverages in C.

SADJUSTEDRESP C
Stores the final adjusted response values in C.

SIWEIGHTS C
Stores the internal weights used to estimate the model parameters in C.

SECOEFFICIENTS C
Stores the standard errors of the estimated coefficients in C.

COEFFICIENTS C
Stores the estimated coefficients in C.

DBETA C
Stores the change in the estimated regression coefficients (delta b) when you delete all observations with a particular factor/covariate pattern in C. Use the delta b values to detect observations with a strong influence on the coefficients.

DSBETA C
Stores the change in the estimated regression coefficients when you delete all observations with a particular factor/covariate pattern based on the standardized Pearson residual in C.

DCHISQUARE C
Stores the change in the chi-square statistic when you delete all observations with a particular factor/covariate pattern in C. Observations that are poorly fit by the model have high delta chi-square values.

DDEVIANCE C
Stores the change in the deviance when you delete all observations with a particular factor/covariate pattern in C.

XMATRIX *M*
Stores the design matrix in M.

XPWXINVERSE *M*
Stores the X'WX inverse matrix of the estimated coefficients in M.

OPTDES: Session command for selecting an optimal design

OPTDES K [C...C]
Select an optimal design by specifying the number of final design points in K and optionally, specifying candidate design columns in C...C.

The purpose of an optimal design is to select design points according to some criteria. Minitab's optimal design capabilities can be used with response surface designs and mixture designs. You can use Select Optimal Design to do the following:

- Select an optimal set of design points.
- Augment (add points to) an existing design.
- Improve the D-optimality of an existing design.
- Evaluate and compare designs.

OPTDES has the following forms:

- OPTDES K: Specifies the number of final design points to select based on D-optimality. For this command form either a factorial model, mixture model, or response surface model must be specified by subcommands FACMODEL, MIXMODEL, or RESMODEL. The constant K must be a positive integer or 0. If K = zero, you must use either the AUGMENT or EVALUATE subcommand.
- OPTDES K, C...C: Specifies the number of final design points and the design columns that contain the candidate set. Constant K must be a positive integer. You must specify at least 2 design columns. All columns must have the same length and must be numeric. With this form of the main command you cannot have the following subcommands: FACMODEL, RESMODEL, MIXMODEL, EVALUATE, BLOCK, GENERATE, EXCHANGE, FEDOROV, SEQUENTIAL.

Note The subcommands DESIGN and COPY replace the obsolete subcommands XMATRIX and CARRY.

FACMODEL *terms*
Species the terms of a factorial model used in selecting D-optimal design. The legal term list is the same as in GLM.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

MIXMODEL *terms*

Specifies the terms of a mixture model used in selecting D-optimal design. The legal term list is the same as in MIXREG.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

RESMODEL *terms*

Specifies the terms of a response surface model used in selecting D-optimal design. The legal term list is the same as in RSREG. When you use RESMODEL, also use INUNIT and LEVELS.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

INUNIT

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

LEVELS K K, ..., K K

Specifies uncoded levels for all continuous factors K K, ..., K K. Enter the low level, then the high level, for each continuous factor that follows RESMODEL. List the pairs of levels in the order of the factors that follow RESMODEL.

AUGMENT C

Augments or improves an existing design indicated by row flags in C. Column C must be numeric and must contain integers. The column length must match the length of the design columns.

Cannot be specified together with EVALUATE, GENERATE, or SEQUENTIAL.

EVALUATE C

Evaluates an existing design indicated by row flags in C. Column C must be numeric and must contain non-negative integers. The column length must match the length of the design columns.

Cannot be specified with OPTDES K C...C, AUGMENT, EXCHANGE, FEDOROV, SEQUENTIAL, GENERATE, INDICATOR.

BLOCK C

Specifies the blocking column. The column can be alpha or numeric with 2–64 distinct values. Column length must match the length of the design columns. Must not be on the term list.

Cannot be specified with OPTDES K C...C.

PROCESS C...C

Specifies process variable columns. Columns must be either on the main line if OPTDES K C...C or on the term list. Columns may be alpha or numeric with exactly 2 levels. With OPTDES K C...C, only numerical process variable columns are allowed. The number of columns must be between 1 and 7. Column lengths must match the length of design columns.

Cannot be specified with AMOUNT or with factorial or response surface models.

AMOUNT C

Specifies the mixture amount column. Column must be either on the main line if OPTDES K C...C or on the term list. The column must be numeric with 2–5 distinct numerical values. The column length must match the length of the design columns.

Cannot be specified with PROCESS or with factorial or response surface models.

PSEUDO

Specifies that calculations are performed using pseudo-components. PSEUDO is ignored in cases of factorial or response surface designs.

GENERATE K [K [K]]

Generates K % of initial points randomly, and adds the rest of points by sequential method. The second K is the number of random trials. The third K is the base for the random data generator. Legal values are as follows: first and second K integer, third K any positive real. 0 < first K <=100, 1 <= second K <= 25.

Cannot be specified with OPTDES K C...C, AUGMENT, EVALUATE, SEQUENT.

SEQUENTIAL

Selects design points by the sequential algorithm for the initial design.

Cannot be specified with OPTDES K C...C, AUGMENT, EVALUATE, GENERATE.

EXCHANGE K

Improves the initial design matrix by adding and dropping K design points one at a time. Legal values are as follows: K is an integer, 1 <= K <= 5.

Cannot be specified with OPTDES K C...C, EVALUATE, FEDOROV.

FEDOROV

Improves the initial design matrix using Fedorov's algorithm.

Cannot be specified with OPTDES K C...C, EVALUATE, EXCHANGE.

INDICATOR C

Stores selection indicators in the specified column.

Cannot be specified with EVALUATE.

DESIGN

Copies the standard order, run order, point type, blocks, and factors to a new worksheet for rows in the optimal design.

COPY C...C

Copies rows in the optimal design to the new worksheet for columns that are not part of the design. For example, you can use COPY to copy a column of covariate measurements.

BRIEF K

Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage.

Value of K	Output that is displayed
1	Minitab displays a summary table for the final design only.
2 (default)	Minitab displays summary tables for the intermediate design and the final design.
3	Displays summary tables and the final design matrix.

DETZERO [K]

Overrides the lower limit for the determinant, which is E-30. Use when you receive the "rank deficient data" error message, but still want to perform the analysis. K is an integer that is greater than or equal to 30. The default value is 30.

Plot Response Surface Designs

FACPLOT: Session command for creating a factorial plot

FACPLOT C

FACPLOT "T|K"

Displays factorial plots for the model associated with the response variable in the specified column. To create this graph, you need to specify a model. To produce an interaction plot, you must have two or more variables.

Displays factorial plots for the model that you specify. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following code: `FACPLOT 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `FACPLOT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

To produce an interaction plot, you must have two or more variables. These plots can illustrate how a response variable relates to one or more variables. You can produce two types of factorial plots.

- Use a main effects plot to compare fitted means or data means across levels of one or more variables.
- Use an interactions plot to assess the effects of one variable at levels of another variable. An interactions plot displays the fitted means for a level of a variable with the level of a second variable held constant.

Although you can use these plots to display the effects, be sure to evaluate significance by looking at the effects in the analysis of variance table.

FACTORS C...C

Specifies the variables to display in both the main effect and interaction plots. You must specify 2 or more variables to generate an interaction plot.

GMAIN

GMAIN displays the main effect plot.

GINT

GINT displays the interactions plot.

FULL

Use FULL to display both interaction plots for each pair of categorical variables, such as Predictor 1 by Predictor 2 and Predictor 2 by Predictor 1. Sometimes you can identify interactions more easily when you examine both plots.

LOWER

Use LOWER to display one version of each interaction plot.

GMODEL (default)

GMODEL display plots for only terms that are in the model and listed in FACTORS. For purposes of a main effect plot, a variable is in the model if the model includes any terms that involve only that variable, e.g. A, A*A, or A*A*A. For purposes of an interaction plot, the interaction of two variables is in the model if the model includes any terms that involve both variables, e.g. A*B, A*B*B.

GALL

GALL displays plots for all terms involving variables on the FACTORS subcommand.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

Mixture Designs

EVDESIGN: Session command for creating an extreme vertices design

EVDESIGN K

Creates an extreme vertices design for K components. In EVDESIGN, Minitab uses an algorithm that generates extreme vertices and their blends up to the specified degrees. You must use these designs when your chosen design space is not an L-simplex design. The presence of both lower and upper bound constraints on the components often create this condition. The goal of an extreme vertices design is to choose design points that adequately cover the design space.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.).

Design

DEGREE K

Specifies the degree of the design in K. For lattice designs, 1 ≤ K ≤ 10. For extreme vertices designs, 1 ≤ K ≤ number of components. The default K = 1.

AUGMENT

Augments (that is, adds points to) the design using axial points. Each axial point is added halfway between a vertex and the center of the design.

By augmenting a design, you can better understand what occurs on the interior of the design, instead of relying only on points on the edges.

AXIAL

Augments the design with axial points halfway between each vertex and centroid. The default is no axial points.

CENTER

Augments the design with center points if no center points exist in the design. The default is no center point.

ALLREP K

Replicates all points in the design K times. The default is K = 1.

REPS K, K...K, K

Specifies the number of replicates for design points of the selected type. In each constant pair (K, K), the first K value indicates the point type (1, 2,...q, 0, –1), and the second K value indicates the number of replicates. The default is no replicates.

Components

TOTAL K [K...K]

TOTAL C

Specifies the mixture design as actual measurements, not proportions, in K or C, with up to 5 different amounts. For example, suppose you measure all the components of a mixture in liters. If the total mixture adds up to 5.2 liters, set TOTAL to 5.2. This generates the settings for components in the design in terms of liters. Or, suppose you want to evaluate plant growth when one gram versus two grams of food are applied. You set TOTAL to 1 2.

K can be any positive number, with up to 5 values. The default value is 1. You can also enter the values in a column (C) and use the column as the argument.

LOWER K...K

LOWER C

Specifies the lower bounds for the q components in K...K, one for each component. The first argument is for the first component, the second for the second component, and so on. The default value is 0. If the sum of all the lower bounds is less than the value of TOTAL, you can specify any non-negative values for LOWER. You can also enter the values in a column (C) and use the column as the argument for LOWER.

When you change the default lower bounds of one component, the achievable upper values on the other components are also adjusted downward. Specifically, the achievable upper bound (U_l) for component l is as follows:

U_l = TOTAL – [L_1 + L_2 + ... + L ($_{l-1}$) + L ($_{l+1}$) + ... + L_q]

Minitab displays both the specified lower bounds and the achievable upper bounds in the output. For example, suppose you have 3 components: TOTAL = 10, L_1 = 0, L_2 = 2, and L_3 = 1. In this case, the achievable upper bounds are U_1 = 7, U_2 = 9, and U_3 = 8.

UPPER K...K

UPPER C

Specifies upper bounds for the components in K...K or C. Must be a positive value for each component. The default is 1 for each value if the units are proportions or pseudo-components. The default is the whole amount if the units are amounts.

PSEUDO and PROPORTIONS

Specifies the scale in which data are represented, including: amounts, proportions, or pseudo-components. Amounts are what you actually measure. By default, Minitab uses amounts to display and store your data. PSEUDO specifies that the data is represented by pseudo-components. PROPORTIONS specifies that the data is represented by proportions.

To get proportions, divide the measurement amounts by the value of TOTAL. When you use PROPORTIONS, Minitab uses proportions to print and store the design. When you do not specify TOTAL, the amount scale and the proportion scale are the same.

You generally use pseudo-components when you specify lower bounds. Pseudo-components, in effect, rescale the constrained data area so that the least allowable amount (the lower bound) of each component is 0, and the greatest allowable amount of a component is 1. This makes the pseudo-component terms of a constrained design (that is, a design with lower bounds) the same as the proportional terms of an unconstrained design.

Mathematically, this translates to the following: Let X_1 be the settings expressed as proportions. Let L_i be the lower bounds expressed as proportions. Let Low = $L_1 + L_2 + ... + L_q$. Then pseudo-component $p_i = (X_i - L_i) / (1 - \text{Low})$.

When you use PSEUDO, then Minitab uses pseudo-components to print and store the design. When the design has no LOWER bounds, pseudo-components and proportions are the same. When you do not use TOTAL or LOWER, the PSEUDO, PROPORTIONS, and amount scales are the same.

You cannot use both PSEUDO and PROPORTIONS in the same command.

In the following example, 2 components show amounts, proportions, and pseudo-components. Suppose the total mixture is 50 ml. Let X1 and X2 be the amount scale. Thus X1 + X2 = 50. Suppose X1 has a LOWER bound of 20. (This means that the UPPER bound of X2 is 50 – 20, or 30.)

```
SCDESIGN 2;
 TOTAL 50;
 LOWER 20  0.
```

The following table shows some points on the three scales.

Amounts		Proportions		Pseudo-components	
X1	**X2**	**X1**	**X2**	**X1**	**X2**
50	0	1.0	0.0	1.0	0.0
20	30	0.4	0.6	0.0	1.0
35	15	0.7	0.3	0.5	0.5

LINEAR K...K

LINEAR C

Specifies linear constraints for the components in K...K or C. The number of values equals the number of components + 2. The first constant is the lower bound. The last constant is the upper bound. One of the bounds can be ' * '. All values must be numeric. The default is no linear constraint.

Process variables

PROC K1 [K2 [K3]]

Specifies the number of process variables, the factorial design, and the fraction id. The default is no process variables. The Ks can equal: 1 ≤ K1 ≤ 7; K2 = 1, 2, 4, 8, 16; 1 ≤ K3 ≤ K2.

LEVELS K...K

Specifies level values for the process variables in K...K. The default is –1 and +1 for each variable. The number of Ks equals twice the number of process variables. K can be numeric or text.

Options

RANDOMIZE [K]

Randomizes the runs in the design matrix. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random data generator and can take any integer as an argument. To always get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the mixture components. Enclose names in quotation marks. If you do not use NAMES, then the names of the components are letters, skipping the letter I: A, B, C...

PNAMES K K, ..., K K

Specifies names for the process variables. Enclose names in quotation marks. If you do not use PNAMES, then the names of the process variables are the letter X followed by sequential numbers: X1, X2, X3...

STOTAL C

Stores one or more totals in C. The default is no storage. STOTAL is incompatible with DESIGN.

SLOWER C

Stores the lower bounds of the components in C. The default is no storage. SLOWER is incompatible with DESIGN.

SUPPER C

Stores the upper bounds of the components in C. The default is no storage. SUPPER is incompatible with DESIGN.

SLINEAR C...C

Stores linear constraints in C...C, one per column. The default is no storage. SLINEAR is incompatible with DESIGN.

Results

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all requested storage.

Value of K	Output that Minitab displays
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the data matrix.

SCDESIGN: Session command for creating a simplex centroid design

SCDESIGN K

Creates a simplex centroid design for K components.

Generates a series of settings for the components in an experiment with a simplex centroid design. A simplex centroid design consists of the following for a mixture with *q* components:

- All points (x1, x2, .., xq) where one component, xi = 1, and the other components are 0. These are called vertex points.
- All points where one component, xi = 1/2, another component, xj = 1/2, and the other components are 0.
- All points where one component, xi = 1/3, another component, xj = 1/3, another component, xk = 1/3, and the other components are 0.
- And so on, up to the stage where all components are 1/q. This last point (where all components are equal) is called the center or centroid of the design.

Simplex centroid designs contain a total of 2q – 1 points.

The constant argument on the command line is an integer from 2 to 7 that specifies the number of components or variables (which is generally called *q*).

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.).

Design

AUGMENT

Augments (or adds points to) the design using the axial points shown below. Each added point is half way between a vertex and the center of the design.

By augmenting a design, you can get a better picture of what happens on the interior of the design, instead of just relying on points on the edges.

AXIAL

Augments the design with axial points half way between each vertex and centroid. The default is no axial points.

ALLREP K

Replicates all points in the design K times. The default is K = 1.

REPS K K, ..., K K

Specifies the number of replicates for design points of selected type. In each constant pair K, K the first value indicates the point type (1, 2, ..q, 0, –1), the second value specifies the number of replicates. The default is no replicates.

Components

TOTAL K [K...K]

TOTAL C

Specifies the mixture design in terms of actual measurements, not proportions. For example, if you measure all the components of your mixture in liters and the total mixture adds up to 5.2 liters, set TOTAL to 5.2 to generate the settings for components in the design in terms of liters. Or, suppose you are testing plant food and would like to evaluate plant growth when one gram versus two grams are applied. You would set TOTAL to 1 2.

The TOTAL argument can be any positive number, with up to five values for total. The default value is 1. You can also enter the values in a column and use the column as the argument.

LOWER K...K

LOWER C

Specifies the lower bounds for the q components. The first argument is for the first component, the second for the second component, and so on. The default is to set the lower bound for each component to 0. You may specify any non-negative values for LOWER provided the sum of all the lower bounds is less than the value of TOTAL. You can also enter the values in a column and use the column as the argument for LOWER.

When you change the default lower bounds of one component, the achievable upper values on the other components are also adjusted downward. Specifically, the achievable upper bound (U i) for component i is

U_i = TOTAL – [L_1 + L_2 + ... + $L_{(i-1)}$ + $L_{(i+1)}$ + ... + L_q]

Minitab displays both the specified lower bounds and achievable upper bounds in the output. As an example, suppose you have 3 components, TOTAL = 10, L_1 = 0, L_2 = 2 and L_3 = 1. Then the achievable upper bounds are U_1 = 7, U_2 = 9 and U_3 = 8.

UPPER K...K

UPPER C

Specifies upper bounds for the components in K...K or C. Must be a positive value for each component. The default is 1 for each value if the units are proportions or pseudo-components. The default is the whole amount if the units are amounts.

PSEUDO and PROPORTIONS

Specifies the scale in which data is represented, including: amounts, proportions, or pseudo-components. Amounts are what you actually measure. By default, Minitab uses amounts to print and store your data.

To get proportions, divide the measurement amounts by the value of TOTAL. When you use PROPORTIONS, Minitab uses proportions to print and store the design. When you do not specify TOTAL, the amount scale and the proportion scale are the same.

You generally use pseudo-components when you specify lower bounds. Pseudo-components, in effect, rescale the constrained data area so the least allowable amount (the lower bound) of each component is 0, and the greatest allowable amount of a component is 1. This makes the pseudo-component terms of a constrained design (that is, a design with lower bounds) the same as the proportional terms of an unconstrained design.

Mathematically, this translates to the following. Let x 1 be the settings expressed as proportions. Let Li be the lower bounds expressed as proportions and let Low = L_1 + L_2 + & + L_q. Then pseudo-component p_i = (x_i – L_i) / (1 – Low).

When you use PSEUDO, then Minitab uses pseudo-components to print and store the design. When the design has no LOWER bounds, pseudo-components and proportions are the same. When you do not use TOTAL or LOWER, the PSEUDO, PROPORTIONS, and amount scales are the same.

You cannot use both PSEUDO and PROPORTIONS in the same command.

Here is an example with 2 components that shows amounts, proportions, and pseudo-components. Suppose the total mixture is 50 ml. Let X1 and X2 be the amount scale. Thus X1 + X2 = 50. Suppose X1 has a LOWER bound of 20 (this means that the UPPER bound of X2 is 50 minus 20, or 30). The session commands for this are as follows.

```
SCDESIGN 2;
 TOTAL 50;
 LOWER 20 0.
```

The following table shows some points on the three scales.

Amounts		Proportions		Pseudo-components	
X1	**X2**	**X1**	**X2**	**X1**	**X2**
50	0	1.0	0.0	1.0	0.0
20	30	0.4	0.6	0.0	1.0
35	15	0.7	0.3	0.5	0.5

Process variables

PROC K1 [K2 [K3]]

Specifies the number of process variables, the factorial design and fraction id. The default is no proc vars 1 <= K1 <= 7; K2 = 1, 2, 4, 8, 16; 1 <= K3 <= K2.

LEVELS K...K

Specifies level values for the process variables. The default is –1 and +1 for each variable. The number of Ks equals twice the number of process variables. K may be numerical or text.

Options

RANDOMIZE [K]

Randomizes the runs in the design matrix. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random data generator and can take any integer as an argument. To always get the same ordering of the rows, create the same design and randomize the runs with the same base. See BASE on page 87 for a discussion.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the mixture components. Enclose names in quotation marks. If you do not use NAMES, then the names of the components are letters, skipping the letter I: A, B, C...

PNAMES K K, ..., K K

Specifies names for the process variables. Enclose names in quotation marks. If you do not use PNAMES, then the names of the process variables are the letter X followed by sequential numbers: X1, X2, X3...

STOTAL C

Stores one or more totals in C. The default is no storage. STOTAL is incompatible with DESIGN.

SLOWER C
Stores the lower bounds of the components in C. The default is no storage. SLOWER is incompatible with DESIGN.

SUPPER C
Stores the upper bounds of the components in C. The default is no storage. SUPPER is incompatible with DESIGN.

Results

BRIEF K
Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output from the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

SLDESIGN: Session command for creating a mixture design

SLDESIGN K [K]
Generates a series of settings for an experiment that uses a simplex lattice design with q components (variables), of degree m. The degree m can be 1–10. For more information, go to Simplex lattice design descriptions on page 1263.

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands SORDER, PTTYPE, and XMATRIX.).

Design

DEGREE K
Specifies the degree of the design in K. For lattice designs, 1 ≤ K ≤ 10. For extreme vertices designs, 1 ≤ K ≤ number of components. The default K = 1.

AUGMENT
Augments (that is, adds points to) the design using axial points. Each axial point is added halfway between a vertex and the center of the design.

By augmenting a design, you can better understand what occurs on the interior of the design, instead of relying only on points on the edges.

AXIAL
Augments the design with axial points halfway between each vertex and centroid. The default is no axial points.

CENTER
Augments the design with center points if no center points exist in the design. The default is no center point.

ALLREP K

Replicates all points in the design K times. The default is K = 1.

REPS K, K...K, K

Specifies the number of replicates for design points of the selected type. In each constant pair (K, K), the first K value indicates the point type (1, 2,...q, 0, –1), and the second K value indicates the number of replicates. The default is no replicates.

Components

TOTAL K [K...K]

TOTAL C

Specifies the total mixture amount in actual measurements in K or C. The TOTAL argument can be any positive number, with up to five values for total. The default value is 1. You can also enter the values in a column and use the column as the argument.

Specifies the mixture design in terms of actual measurements, not proportions. For example, if you measure all the components of your mixture in liters and the total mixture adds up to 5.2 liters, set TOTAL to 5.2 to generate the settings for components in the design in terms of liters. Or, suppose you are testing plant food and want to evaluate plant growth when one gram versus two grams are applied. You would set TOTAL to 1 2.

LOWER K...K

LOWER C

Specifies the lower bounds for the q components in K...K, one for each component, or in C.

The first argument is for the first component, the second for the second component, and so on. The default is to set the lower bound for each component to 0. You may specify any non-negative values for LOWER provided the sum of all the lower bounds is less than the value of TOTAL. You can also enter the values in a column and use the column as the argument for LOWER.

When you change the default lower bounds of one component, the achievable upper values on the other components are also adjusted downward. Specifically, the achievable upper bound (U_i) for component i is

$U_i = \text{TOTAL} - [L_1 + L_2 + ... + L_{(i-1)} + L_{(i+1)} + ... + L_q]$

Minitab displays both the specified lower bounds and achievable upper bounds in the output. For example, suppose you have 3 components, TOTAL = 10, $L_1 = 0$, $L_2 = 2$ and $L_3 = 1$. Then the achievable upper bounds are $U_1 = 7$, $U_2 = 9$ and $U_3 = 8$.

UPPER K...K

UPPER C

Specifies upper bounds for the components in K...K or C. Must be a positive value for each component. The default is 1 for each value if the units are proportions or pseudo-components. The default is the whole amount if the units are amounts.

PSEUDO and PROPORTIONS

Specifies the scale in which data are represented, including: amounts, proportions, or pseudo-components. Amounts are what you actually measure. By default, Minitab uses amounts to display and store your data. PSEUDO specifies that the data is represented by pseudo-components. PROPORTIONS specifies that the data is represented by proportions.

To get proportions, divide the measurement amounts by the value of TOTAL. When you use PROPORTIONS, Minitab uses proportions to print and store the design. When you do not specify TOTAL, the amount scale and the proportion scale are the same.

You generally use pseudo-components when you specify lower bounds. Pseudo-components, in effect, rescale the constrained data area so that the least allowable amount (the lower bound) of each component is 0, and the greatest allowable amount of a component is 1. This makes the pseudo-component terms of a constrained design (that is, a design with lower bounds) the same as the proportional terms of an unconstrained design.

Mathematically, this translates to the following: Let X_1 be the settings expressed as proportions. Let L_i be the lower bounds expressed as proportions. Let Low = $L_1 + L_2 + ... + L_q$. Then pseudo-component $p_i = (X_i - L_i) / (1 - \text{Low})$.

When you use PSEUDO, then Minitab uses pseudo-components to print and store the design. When the design has no LOWER bounds, pseudo-components and proportions are the same. When you do not use TOTAL or LOWER, the PSEUDO, PROPORTIONS, and amount scales are the same.

You cannot use both PSEUDO and PROPORTIONS in the same command.

In the following example, 2 components show amounts, proportions, and pseudo-components. Suppose the total mixture is 50 ml. Let X1 and X2 be the amount scale. Thus X1 + X2 = 50. Suppose X1 has a LOWER bound of 20. (This means that the UPPER bound of X2 is 50 – 20, or 30.)

```
SLDESIGN 2;
 TOTAL 50;
 LOWER 20  0.
```

The following table shows some points on the three scales.

Amounts		Proportions		Pseudo-components	
X1	**X2**	**X1**	**X2**	**X1**	**X2**
50	0	1.0	0.0	1.0	0.0
20	30	0.4	0.6	0.0	1.0
35	15	0.7	0.3	0.5	0.5

Process variables

PROC K1 [K2 [K3]]

Specifies the number of process variables, the factorial design, and the fraction id. The default is no process variables. The Ks can equal: 1 ≤ K1 ≤ 7; K2 = 1, 2, 4, 8, 16; 1 ≤ K3 ≤ K2.

LEVELS K...K

Specifies level values for the process variables in K...K. The default is –1 and +1 for each variable. The number of Ks equals twice the number of process variables. K can be numeric or text.

Options

RANDOMIZE [K]

Randomizes the runs in the design matrix. When you use RANDOMIZE, the standard order and the run order are different and Minitab stores the design in run order.

K is the base for the random data generator and can take any integer as an argument. To always get the same ordering of the rows, create the same design and randomize the runs with the same base. For more information, go to BASE on page 87.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the mixture components. Enclose names in quotation marks. If you do not use NAMES, then the names of the components are letters, skipping the letter I: A, B, C...

PNAMES K K, ..., K K

Specifies names for the process variables. Enclose names in quotation marks. If you do not use PNAMES, then the names of the process variables are the letter X followed by sequential numbers: X1, X2, X3...

STOTAL C

Stores one or more totals in C. The default is no storage. STOTAL is incompatible with DESIGN.

SLOWER C

Stores the lower bounds of the components in C. The default is no storage. SLOWER is incompatible with DESIGN.

SUPPER C

Stores the upper bounds of the components in C. The default is no storage. SUPPER is incompatible with DESIGN.

Results

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output from the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

MIXREG: Session command for analyzing a mixture design

MIXREG C...C = C...C

Analyzes (or fit) Scheffe mixture models, including simplex centroid and simplex lattice designs. You have a choice of four standard models: linear, quadratic (default), special cubic, special quadratic, full cubic, and full quadratic.

The main command contains a list of the response variables before the equal sign and a list of the components after the equal sign. If the design is blocked, give the blocking variable with the BLOCKS subcommand.

You can have up to 25 responses before the equal sign. For information on how Minitab handles missing data, go to Missing values in factorial, response surface, and mixture designs on page 1254.

To use commands on the Minitab menu that use model information, such as the response optimizer, use the subcommand DESIGN.

PROC C...C

Specifies the columns that contain process variables. The number of columns must be from 1 to 7. The process variables must be specified in term list.

AMOUNT C

Specifies the column that contains the mixture amounts. This column should not be specified in MIXREG on the term list.

INUNIT K

Specifies the unit of the components.

Value of K	Unit of the components
0	Pseudo-component units
1	Proportion units
2	Amount units

PSEUDO

Specifies that the analysis is performed in pseudo-components.

Stepwise

Specifies a model selection procedure. STEPWISE, FORWARD, BACKWARD are the three subcommands that correspond to stepwise, forward, and backward model selection procedures.

You can specify one model selection procedure at a time.

STEPWISE

If you specify STEPWISE, then the FORCE, ENTER, AENTER, AREMOVE, and BEST subcommands are valid.

FORWARD

If you specify FORWARD, then the FORCE, AENTER, and BEST subcommands are valid.

BACKWARD

If you specify BACKWARD, then the FORCE, AREMOVE, and BEST subcommands are valid.

FORCE *terms*

Specifies the terms to be forced in the model. Main component terms are always present. Terms selected must be a subset of the model terms given on MIXREG.

ENTER *terms*

Specifies the terms that are contained in a starting model for stepwise selection procedure. The term entered here must be one of the model terms specified on MIXREG

Only the STEPWISE subcommand uses ENTER.

AENTER K

Specifies the alpha level at which a term is entered to the model. K is a constant between 0 and 1 (including 0 and 1). The default value is 0.15 for STEPWISE, and 0.25 for FORWARD.

For stepwise selection, K for AREMOVE must be greater than or equal to K for AENTER. If only AREMOVE is specified by a user, then AENTER = AREMOVE, vice versa.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. K is a constant between 0 and 1 (including 0 and 1). The default value is 0.15 for STEPWISE, and 0.1 for BACKWARD. The FORWARD subcommand doesn't use this input.

For stepwise selection, K for AREMOVE must be greater than or equal to K for AENTER. If only AREMOVE is specified by a user, then AENTER = AREMOVE, vice versa.

BEST K

Specifies the best K alternative terms to show in the output. K is an integer from 1 to 10.

If BEST K is not entered, no alternative term will appear in the output. The default is no alternative term.

Terms

LOWER K...K

If you do not use LOWER, the lower bound of each component is the minimum value in the column of the component. Specify as many bounds as you specify components. Specify 0 for components that do not have lower bounds.

BLOCKS C

Specifies a blocking variable column when you specify your model explicitly on MIXREG. That is, you are not using a model that you have stored with another command. The blocking variable column can be numeric or text, and can contain any value. The levels do not need to be in any special order. When the blocking column is numeric, the smallest value is block 1, the second smallest value is block 2, and so on.

Prediction

PREDICT *E...E*

Computes the fitted response values for given values of the components. PREDICT displays a table that contains the fitted Ys, standard errors of the fitted Ys, a 95% confidence interval, and a 95% prediction interval. E...E can be a list of constants or a list of columns, or a mixture of constants and columns. Component columns or constants should be listed first, followed by process variables, if specified, and the amount variable, if specified.

PBLOCK *E*

Specifies the block of new design points if blocks are included in your design.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PSFITS C

Stores the fits in C.

PSEFIT C

Stores the estimated standard errors of the fits in C.

CLIMITS C...C

Stores the lower and upper confidence limits in C...C.

PLIMITS C...C

Stores the lower and upper prediction limits in C...C.

Graphs

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data, or versus the order specified in C. The row number for each data point is shown on the x-axis (for example, 1 2 3 4...n).

GFOURPACK [C]

Displays four residual plots—a histogram of the residuals, a normal probability plot of the residuals, the residuals versus fitted the values, and the residuals versus the order of the data, or versus the order specified in C—in one graph.

GVARIABLE C...C

Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Results

TOLERANCE K

Specifies the tolerance level for collinearity and constant check. By default, K = 4 * 2.22e-016

BRIEF K

Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the table of coefficients, s, R-sq, R-sq(adj), the analysis of variance table, the unusual values in the table of fits and residuals.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but also displays the full table of fits and residuals.

ORDER C
Specifies the row order of the table of fits and residuals that is displayed in the output. Specify the StdOrder column in C.

Storage

FITS C...C
Stores the fitted values in C...C.

RESIDUALS C...C
Stores the residuals (response – fit) in C...C.

SRESIDUALS C...C
Stores the standardized residuals in C...C.

TRESIDUALS C...C
Stores the deleted Studentized residuals in C...C.

COEFFICIENTS C...C
Stores the coefficients, using one column for each response. These are the same coefficients that are displayed in the output. If MIXREG removes some terms because the data cannot support them, then the removed terms do not appear on the output.

MODEL C...C
Requests storage of model term codes. The number of columns matches the number of responses in C...C.

XMATRIX *M*
Stores the design matrix. If the design was blocked into k blocks, then there are (k – 1) columns for block dummy variables. MIXREG uses the same method of coding blocks as GLM. The block dummy variables are followed by one column for each component. When the model has 2-way interactions, the design matrix contains a column for each interaction. The column for a 2-way interaction is the product of the corresponding two components. When the model has 3-way interactions, the design matrix contains a column for each interaction. The column for a 3-way interaction is the product of the corresponding three components. When the model has additional terms for a full cubic, then the design matrix contains one column for each additional term. For example, the column for a term such as X1 * X3 * (X1 – X3), with X1 in C1 and X3 in C3, is calculated using the equation C1 * C3 * (C1 – C3).

Because MIXREG allows no constant in the model, the design matrix does not contain a column of 1s.

When MIXREG removes terms because the data cannot support them, the design matrix does not contain the removed terms. The columns of the stored matrix match the coefficients that are displayed and/or stored.

HI C
Stores the leverage values.

COOKD C...C
Stores Cook's distance values, using one column for each response.

DFITS C...C
Stores the DFITS (also called DFFITS), using one column for each response.

DESIGN

Stores model information to enable items on the menu that use the model, such as the response optimizer. To use session commands to run the analyses that use the model, store the model information in a column with MODEL.

OPTDES: Session command for selecting an optimal design

OPTDES K [C...C]

Select an optimal design by specifying the number of final design points in K and optionally, specifying candidate design columns in C...C.

The purpose of an optimal design is to select design points according to some criteria. Minitab's optimal design capabilities can be used with response surface designs and mixture designs. You can use Select Optimal Design to do the following:

- Select an optimal set of design points.
- Augment (add points to) an existing design.
- Improve the D-optimality of an existing design.
- Evaluate and compare designs.

OPTDES has the following forms:

- OPTDES K: Specifies the number of final design points to select based on D-optimality. For this command form either a factorial model, mixture model, or response surface model must be specified by subcommands FACMODEL, MIXMODEL, or RESMODEL. The constant K must be a positive integer or 0. If K = zero, you must use either the AUGMENT or EVALUATE subcommand.
- OPTDES K, C...C: Specifies the number of final design points and the design columns that contain the candidate set. Constant K must be a positive integer. You must specify at least 2 design columns. All columns must have the same length and must be numeric. With this form of the main command you cannot have the following subcommands: FACMODEL, RESMODEL, MIXMODEL, EVALUATE, BLOCK, GENERATE, EXCHANGE, FEDOROV, SEQUENTIAL.

Note The subcommands DESIGN and COPY replace the obsolete subcommands XMATRIX and CARRY.

FACMODEL *terms*

Species the terms of a factorial model used in selecting D-optimal design. The legal term list is the same as in GLM.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

MIXMODEL *terms*

Specifies the terms of a mixture model used in selecting D-optimal design. The legal term list is the same as in MIXREG.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

RESMODEL *terms*

Specifies the terms of a response surface model used in selecting D-optimal design. The legal term list is the same as in RSREG. When you use RESMODEL, also use INUNIT and LEVELS.

FACMODEL, RESMODEL, and MIXMODEL are mutually exclusive and one of them must be the first subcommand.

INUNIT

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

LEVELS K K, ..., K K

Specifies uncoded levels for all continuous factors K K, ..., K K. Enter the low level, then the high level, for each continuous factor that follows RESMODEL. List the pairs of levels in the order of the factors that follow RESMODEL.

AUGMENT C

Augments or improves an existing design indicated by row flags in C. Column C must be numeric and must contain integers. The column length must match the length of the design columns.

Cannot be specified together with EVALUATE, GENERATE, or SEQUENTIAL.

EVALUATE C

Evaluates an existing design indicated by row flags in C. Column C must be numeric and must contain non-negative integers. The column length must match the length of the design columns.

Cannot be specified with OPTDES K C...C, AUGMENT, EXCHANGE, FEDOROV, SEQUENTIAL, GENERATE, INDICATOR.

BLOCK C

Specifies the blocking column. The column can be alpha or numeric with 2–64 distinct values. Column length must match the length of the design columns. Must not be on the term list.

Cannot be specified with OPTDES K C...C.

PROCESS C...C

Specifies process variable columns. Columns must be either on the main line if OPTDES K C...C or on the term list. Columns may be alpha or numeric with exactly 2 levels. With OPTDES K C...C, only numerical process variable columns are allowed. The number of columns must be between 1 and 7. Column lengths must match the length of design columns.

Cannot be specified with AMOUNT or with factorial or response surface models.

AMOUNT C

Specifies the mixture amount column. Column must be either on the main line if OPTDES K C...C or on the term list. The column must be numeric with 2–5 distinct numerical values. The column length must match the length of the design columns.

Cannot be specified with PROCESS or with factorial or response surface models.

PSEUDO

Specifies that calculations are performed using pseudo-components. PSEUDO is ignored in cases of factorial or response surface designs.

GENERATE K [K [K]]

Generates K % of initial points randomly, and adds the rest of points by sequential method. The second K is the number of random trials. The third K is the base for the random data generator. Legal values are as follows: first and second K integer, third K any positive real. 0 < first K <=100, 1 <= second K <= 25.

Cannot be specified with OPTDES K C...C, AUGMENT, EVALUATE, SEQUENT.

SEQUENTIAL

Selects design points by the sequential algorithm for the initial design.

Cannot be specified with OPTDES K C...C, AUGMENT, EVALUATE, GENERATE.

EXCHANGE K

Improves the initial design matrix by adding and dropping K design points one at a time. Legal values are as follows: K is an integer, 1 <= K <= 5.

Cannot be specified with OPTDES K C...C, EVALUATE, FEDOROV.

FEDOROV

Improves the initial design matrix using Fedorov's algorithm.

Cannot be specified with OPTDES K C...C, EVALUATE, EXCHANGE.

INDICATOR C

Stores selection indicators in the specified column.

Cannot be specified with EVALUATE.

DESIGN

Copies the standard order, run order, point type, blocks, and factors to a new worksheet for rows in the optimal design.

COPY C...C

Copies rows in the optimal design to the new worksheet for columns that are not part of the design. For example, you can use COPY to copy a column of covariate measurements.

BRIEF K

Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays a summary table for the final design only.
2 (default)	Minitab displays summary tables for the intermediate design and the final design.
3	Displays summary tables and the final design matrix.

DETZERO [K]

Overrides the lower limit for the determinant, which is E-30. Use when you receive the "rank deficient data" error message, but still want to perform the analysis. K is an integer that is greater than or equal to 30. The default value is 30.

Plot Mixture Designs

SIMPLEX: Session command for creating a simplex design plot

SIMPLEX C...C

Creates a simplex design plot to visualize the mixture design space (or a slice of the design space, if you have more than three components). Minitab plots the design points on triangular axes. You can plot components, components and process variables, or components and an amount variable.

For information about graphics options for SIMPLEX, go to Graphics options for SIMPLEX on page 1237.

SINGLE C C C

Specifies three components for a single plot in C C C. Specified columns must be on the main command list.

ALLTRIPLETS

Requests multiple plots to show all component triplets (for only 3–5 components).

PANEL

Displays plots in one graph.

SEPARATE

Displays each graph in a separate graph.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

PROCESS C...C

Specifies process variable columns in C...C. The length of the columns must match component columns. Columns may be numeric or alpha and must contain 2 levels. The maximum number of process variables is 7.

AMOUNT C

Specifies the mixture amount column in C. The length of the column must match the component columns. The column must be numeric and must contain 2–5 distinct positive values. Must have a Total subcommand with either K = 1, or multiple Ks must match the values in C.

LABEL [C]

Specifies the labels of the design points in C. If you omit C, then label points by number of replicates. The length of the column must match component columns.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

Holds components that are not on the plot to values that you specify in K...K. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

FHOLD C

Holds process variables to values that you specify. Sets the levels of process variables at a combination of –1 and 1. K must equal the number of process variables in the model. Only design points that match these values are plotted. By default, all process variables are held at –1.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

SGRID [K]

Displays grid lines at 0.1 intervals or at intervals specified in K.

SPSIMP: Session command for creating a simplex design plot

SPSIMP C...C

Creates three types of special simplex design plots:

- Simplex plot with four components
- Simplex plot with process variables
- Simplex plot with amount variable

C...C must include all components in the design matrix.

Enables you to visualize the mixture design space (or a slice of the design space if you have more than three components). Minitab plots the design points on triangular axes. You can plot components, components and process variables, or components and an amount variable.

For information on graphics options for SPSIMP, go to Graphics options for SPSIMP on page 1240.

TYPE K

Determines which of the three simplex plots is created in K.

Value of K	Plot that is created	When to use
1	Four component matrix plot	The model contains at least four components.
2	Contour plot with process variables	The model contains process variables.
3	Contour plot with mixture amounts	The amount variable is specified in the model.

SINGLE C C C [C]

Specifies components in C C C [C] on the three or four axes of a single plot.

Three components must be specified if the plot type is 2 or 3. Four components must be specified if the plot type is 1. Columns must be listed in the main command. If this command isn't specified, then Minitab uses the first 3 or 4 components listed in SPSIMP according to the type of plot that you specify in TYPE.

ALLTRIPLETS

Displays multiple plots to show all component triplets (only for 3–5 components).

PANEL

Displays the plots in one graph.

SEPARATE

Displays the plots in separate graphs.

TOTAL K [K...K]

TOTAL C

Specifies 1 to 5 total values. Single K must equal the common row sum of components. Multiple Ks are accepted only if Amount column specified. The total subcommand must be specified if the Amount columns specified either with a single K = 1 or with multiple K values matching the values in column Amount.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

PROCESS C...C

Specifies that the process variables are in C...C. The length of the columns must match component columns. Columns can be numeric or text and must contain 2 levels. The maximum number of process variables is 7.

AMOUNT C

Specifies the mixture amount in C. The length of the column must match the component columns. The column must be numeric, and contain 2–5 distinct positive values. Must have Total subcommand with either K = 1, or multiple Ks must match values in C.

LABEL [C]

Specifies labels for design points are in C. If you omit C, then you should label the points by the number of replicates. The length of the column must match the component columns.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

CHOLD C

Holds components that are not on the plot to values specified in K...K or C. The number of constants or the length of the column must match the number of components. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

FHOLD C

Holds process variables to values that you specify. Sets the levels of process variables at a combination of –1 and 1. K must equal the number of process variables in the model. Only design points that match these values are plotted. By default, all process variables are held at –1.

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines at 0.2 intervals or at intervals specified in K. All K must be 0 < K < 1. The default is K = 0.2.

%SIMPROC: Session command for creating a simplex plot

%SIMPROC C...C

Creates a simplex plot of components in C...C. %SIMPROC is used for certain simplex plots when all level combinations are plotted and there are process variables in the model. A simplex plot helps you visualize the mixture design space. Minitab plots the design points on triangular axes.

SINGLE C C C

Specifies the three components on the three axes in C C C. Columns must be listed in main command.

TOTAL K

Specifies one total value in K. K must equal the common row sum of components.

LOWER K...K

Specifies the lower bound of components in K...K. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL.

UPPER K...K

Specifies the upper bound of components in K...K. The number of K's equals the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL.

LN1 K...K, LN2 K...K, ... LN10 K...K

Specifies the linear constraints in K...K. The first linear constraint is in LN1, the second linear constraint is in LN2, and so on. You can have a maximum of 10 linear constraints.

Each subcommand specifies a linear constraint on components in mixture models. The first K represents the lower bound of a linear combination of the components, and the last K the upper bound of the linear combination. The Ks in between represent the coefficients of the components.

For example, let A, B, and C be three components in a mixture design. A linear constraint on A, B, and C is represented by:

0.5 <= 3A+2B+C<=0.9

Then, the LN1 subcommand is specified as:

```
LN1 0.5 3 2 1 0.9.
```

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PROCESS C C C

Specifies process variable columns in C C C. The length of the columns must match component columns. Columns may be numeric or text and must contain 2 levels. The number of process variables must equal 3.

CHOLD K...K

Specifies the hold values in K...K for components that are not on plot. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted.

CMIN K...K

Specifies the minimum axis values of all components, including those not on the plot, in K...K. Values may be in constants. The number of constants must match the number of components.

SGRID K

Specifies that grid lines are drawn on each axis at intervals specified in K.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

TRACE: Session command for creating a response trace plot

TRACE C C

Displays a response trace plot for the response in the first C and the model column in the second C, which is stored using MIXREG and contains the model information.

A response trace plot (also called component effects plot) shows the effect of each component on the response. Several response traces, which are a series of predictions from the fitted model, are plotted along a component direction. The trace curves show the effect of changing the corresponding component along an imaginary line (direction) connecting the reference blend to the vertex.

Each component in the mixture has a corresponding trace direction. The points along a trace direction of a component are connected thereby producing as many curves as there are components in the mixture.

Specify a single response column, and the column where the model is stored. The model column contains the number and ID of components, number and ID of process variables, ID of blocking column, ID of mixture amount column, term codes and coefficients. The response column length must match the length of component columns.

TOTAL K [K...K]

TOTAL C

Specifies the total amount in K, or in C, from 1–5 values.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The column must be numeric.

The number of values equals the number of components specified in the model column.

All values must be 0 <= K < TOTAL and SUM(K) < TOTAL (TOTAL is row sum of components).

UPPER K...K

UPPER C

Specifies the upper bound of the components. The column must be numeric.

The number of values equals the number of components specified in the model column.

All values must be 0 < K <= TOTAL and SUM(K) > TOTAL (TOTAL is row sum of comps).

CLINEAR K...K

CLINEAR C

Specifies a linear constraint on the components. There can be up to 10 such subcommands. The column must be numeric.

The number of values equals the number of components + 2. The first value, the lower bound, must be smaller than last value, the upper bound. One of the bounds might be missing, i.e., set to ' * '.

Note You can also use LINEAR to specify a linear constraint on the components.

PIEPEL

Specifies the plot uses Piepel directions.

REFERENCE K

REFERENCE K...K

Specifies the row ID of the reference blend or the coordinates of the reference blend. If you do not specify XREFERENCE, then the default is to use the centroid of vertices.

XREFERENCE K specifies row id of the reference blend; 0 < K <= number of design points (length of design columns).

XREFERENCE K...K specifies coordinates of the reference blend. The number of Ks equals the number of components. Each K must be between the corresponding lower and upper bounds.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

Sets process variable levels to specified values.

The number of Ks must equal the number of process variables. Constants must be between -1 and 1.

AHOLD K

Sets mixture amount to a specified value. The specified value must be between min and max of IPAMNT column.

MAXIMUM K K

Specifies the maximum value of an axis. The first K specifies the x- (use 1) or y- (use 2) axis and the second K specifies the maximum value. You can also use MAX to specify the maximum value of an axis.

MINIMUM K K

Specifies the minimum value of an axis. The first K specifies the x- (use 1) or y- (use 2) axis and the second K specifies the minimum value. You can also use MIN to specify the minimum value of an axis.

CURVE

Specifies different type and color for trace curves.

TYPE K...K
Use TYPE to specify the line type of trace curves. The default is solid lines.

COLOR K...K
Use COLOR to specify the line color of trace curves. The default is different colors.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MIXCONTOUR: Session command for creating a contour plot

MIXCONTOUR C C
Displays a contour plot for the response in the first C and the model column in the second C, which is stored using MIXREG and contains the model information.

The model column contains the number and the ID of the components, the number and the ID of the process variables, the ID of the blocking column, the ID of the mixture amount column, the term codes, and the coefficients. The response column and the component columns must have the same number of rows.

For more information about graphics options for MIXCONTOUR, go to Graphics options for MIXCONTOUR on page 1231.

SINGLE C C [C]
Specifies three components in C C C or two numerical process variables in C C for a single plot. You must specify the components or process variables that are on the list in the model column.

ALLTRIPLETS
Specifies that multiple plots show all component triplets (only for 3–5 components).

PANEL
Displays all the plots in one graph.

SEPARATE
Displays each plot on a separate page.

ALLPAIRS
Specifies that multiple plots show all pairs of numerical process variables.

PANEL
Displays all the plots in one graph.

SEPARATE
Displays each plot on a separate page.

NLEVELS K
Specifies the number of contour levels in K. K must be 2–15.

LEVELS K...K
Specifies the values of contour levels in K. K must be 2–15.

NMESH K K
Specifies number of mesh points. The first constant specifies the direction for the number of mesh points:

- 1 = Horizontal

- 2 = Vertical

To specify both directions, issue NMESH twice.

The second constant specifies the number of mesh points. The second K must be 2–101.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

Holds components that are not on the plot to values that you specify in K...K. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

FHOLD C

Holds process variables to values that you specify. Sets the levels of process variables at a combination of –1 and 1. K must equal the number of process variables in the model. Only design points that match these values are plotted. By default, all process variables are held at –1.

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines on each axis at intervals specified in K. All K must be $0 < K < 1$. The default is $K = 0.1$.

SPCONT: Session command for creating a contour design plot

SPCONT C C

Creates three types of special contour design plots:

- Contour plot with four components
- Contour plot with process variables
- Contour plot with an amount variable

The first C is the response, and the second C contains the model information.

Enables you to visualize the mixture design space (or a slice of the design space if you have more than three components). Minitab plots the design points on triangular axes. You can plot components, components and process variables, or components and an amount variable.

For information on graphics options that are available for SPCONT, go to Graphics options for SPCONT on page 1238.

TYPE K

Determines which of the three simplex plots is created in K.

Value of K	Plot that is created	When to use the plot
1	Four component matrix plot	The model contains at least four components.
2	Contour plot with process variables	The model contains process variables.
3	Contour plot with mixture amounts	The amount variable is specified in the model.

SINGLE C C C [C]

Specifies components in C C C [C] on the three or four axes of a single plot.

Three components must be specified if plot type is 2 or 3. Four components must be specified if plot type is 1. Columns must be listed in the model column. If this command isn't specified, Minitab uses the first 3 or 4 components listed in SPCONT according to the type of plot you specified in TYPE.

ALLTRIPLETS

Displays multiple plots to show all component triplets (only for 3–5 components).

PANEL

Displays the plots in one graph.

SEPARATE

Displays the plots in separate graphs.

NLEVELS K

Specifies the number of contour levels. K must be between 2 and 15.

LEVELS K...K

Specifies the values of contour levels in K...K. K must be between 2 and 15.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

PROCESS C...C

Specifies that the process variables are in C...C. The length of the columns must match component columns. Columns can be numeric or alpha and must contain 2 levels. The maximum number of process variables is 7.

AMOUNT C

Specifies the mixture amount in C. The length of the column must match the component columns. The column must be numeric, and contain 2–5 distinct positive values. Must have Total subcommand with either K = 1, or multiple Ks must match values in C.

LABEL [C]

Specifies labels for design points are in C. If you omit C, then you should label the points by the number of replicates. The length of the column must match the component columns.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

CHOLD C

Holds components that are not on the plot to values specified in K...K or C. The number of constants or the length of the column must match the number of components. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

FHOLD C

Holds process variables to values that you specify. Sets the levels of process variables at a combination of –1 and 1. K must equal the number of process variables in the model. Only design points that match these values are plotted. By default, all process variables are held at –1.

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines at 0.2 intervals or at intervals specified in K. All K must be $0 < K < 1$. The default is K = 0.2.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MIXSURFACE: Session command for creating a surface plot

MIXSURFACE C C

Displays surface plot for response in C and the model column in C, which is stored using MIXREG and contains the model information.

Specifies a single response column and the column where the model is stored. The model column contains the number and ID of components, number and ID of process variables, ID of blocking column, ID of mixture amount column, term codes, and coefficients. The length of the response column must match that of component columns.

For information graphics options for MIXSURFACE, go to Graphics options for MIXSURFACE on page 1233.

SINGLE C C [C]

Specifies three components in CCC or two numerical process variables in C C for a single plot. The components or process variables must be specified on the list in the model column.

ALLTRIPLETS

Displays multiple plots to show all component triplets (only for 3–5 components).

PANEL

Displays plots in one graph.

SEPARATE

Displays each graph on a separate page.

ALLPAIRS

Displays multiple plots to show all numerical process variable pairs.

PANEL

Displays plots in one graph.

SEPARATE

Displays each graph on a separate page.

NMESH K K

Specifies number of mesh points. The first constant specifies the direction for the number of mesh points:

- 1 = Horizontal
- 2 = Vertical

To specify both directions, issue NMESH twice.

The second constant specifies the number of mesh points. The second K must be from 2 to 101.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

Holds components that are not on the plot to values that you specify in K...K. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FHOLD K...K

FHOLD C

Holds process variables to values that you specify. Sets the levels of process variables at a combination of –1 and 1. K must equal the number of process variables in the model. Only design points that match these values are plotted. By default, all process variables are held at –1.

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines on each axis at specified intervals. All K must be 0 < K < 1. The default is K = 0.1.

MIXOVER: Session command for creating an overlaid contour plot

MIXOVER

Draws an overlaid contour plot for multiple (up to 10) responses of mixture models.

Contour plots for mixture designs show how response variables are related to three continuous design variables, while holding the other variables in a model at certain settings. In a contour plot, the fitted response model is viewed as a 2-dimensional surface where all points that have the same fitted value are connected to produce contour lines of constants. Contour plots are useful for establishing operating conditions that produce desirable response values.

Applications that involve multiple responses present a different challenge than single-response experiments. Because each response is important in determining the quality of the product, you need to consider the responses simultaneously. Optimal settings for input variables for one response might be far from optimal for another response. Overlaid contour plots allow you to visually identify an area of compromise among the various responses.

For more information about graphics options for MIXOVER, go to Graphics options for MIXOVER on page 1233.

RESPONSE C...C

Specifies a set of up to 10 response variables for the experiment in C...C.

MODEL C...C

Specifies a set of model columns (one for each response), which contain design and model information in C...C. Each model column is stored by the MODEL subcommand of MIXREG.

SINGLE C C [C]

Specifies three components or two numerical process variables for a single plot. You must specify the components or process variables on the list in the model column.

LEVELS K...K

Specifies the values of contour levels in K...K. K must be between 2 and 15.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

Holds components that are not on the plot to values that you specify in K...K. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines on each axis at specified intervals. All Ks must be 0 < K < 1. By default, K is 0.2.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

FFMAIN: Superseded by FACPLOT except for mixture designs

Note FFMAIN is superseded by FACPLOT on page 277, except for mixture designs.

FFMAIN C...C

Creates a main effects plot for 2-level factorial designs, for factors in C...C.

FFMAIN plots the means at the two levels of each factor and connects them with a line. FFMAIN plots any center points you have, using a different symbol than that used for the factorial points. A reference line is drawn through the plot to represent the overall mean (also called grand mean) of the response data.

Specify up to 50 factors. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1. The next entry that is different is level 2. All factors must have 2 levels, and can have center points.

RESPONSE C (required)

Plots the response data in C. If you specify RESPONSE without FMEANS, then Minitab plots the raw response data. The main effects are calculated based on the means of the data. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

With RESPONSE, you can specify up to 50 factors.

FMEANS C...C

Plots the least squares or fitted means in C...C. If you specify RESPONSE with FMEANS, then Minitab plots the predicted values from the model. The main effects are calculated based on the fitted means. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

With FMEANS, you can specify up to 10 factors.

YMIN K

Specifies the minimum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

YMAX K

Specifies the maximum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

CTPT K

Specifies the center point column in K. Minitab considers a value of 0 as a center point.

LEVELS K K, ..., K K

Specifies the levels for the uncoded units for each factor in K K, ..., K K. Use LEVELS to change the default order of levels that Minitab uses.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, which indicates that the factors are process variables.

FFINT: Superseded by FACPLOT except for mixture designs

Note FFINT is superseded by FACPLOT on page 277, except for mixture designs.

FFINT C...C
Creates interaction plots for factors in C...C.

Creates an interaction plot for 2-level factorial designs. If you enter 2 factors on the command line, then Minitab displays a single interaction plot. If you enter more than 2 factors, then Minitab displays a matrix of interaction plots. You can use factors from designs that are generated by FFDESIGN and PBDESIGN.

You can specify up to 15 columns that contain the factors. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on. All factors must have 2 levels, and can have center points.

RESPONSE C (required)
Plots the response data in C. If you specify RESPONSE without FMEANS, then Minitab plots the raw response data. The interaction effects are calculated based on the means of the data. If you plot process variables from a mixture design, then Minitab fits only the raw response data, not the fitted means.

You can specify up to 15 factors.

FMEANS C...C

Plots the least squares or fitted means in C...C. If you specify RESPONSE with FMEANS, then Minitab plots the predicted values from the model. The interaction effects are calculated based on the fitted means. If you plot process variables from a mixture design, Minitab fits only the raw response data, not the fitted means.

FULL

Displays the full interaction matrix when you enter more than 2 factors. By default, Minitab displays the half matrix. The half matrix displays the plots in the upper right portion of the matrix. In the full matrix, the transpose of each plot in the upper right is displayed in the lower left portion of the matrix. The full matrix takes longer to display than the half matrix.

YMIN K

Specifies the minimum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

YMAX K

Specifies the maximum values on the y-axis of every plot that is displayed by the current command in K. Making the minimum and maximum values the same is useful when you compare several plots of related data.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

CTPT C

Specifies the center point column in C. Minitab considers a value of 0 as a center point.

LEVELS K K, ..., K K

Specifies the levels for the uncoded units for each factor in K K, ..., K K. Use LEVELS to change the default order of levels that Minitab uses.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, which indicates that the factors are process variables.

FFCUBE: Session command for creating a cube plot

FFCUBE C...C

Use FFCUBE to draw a cube plot that does not show mean response values. With a mixture design, you can use FFCUBE to draw a cube plot that shows mean response values. For all other models, such as the models from FFAC or REGR, use MFFCUBE on page 337 to draw a cube plot that shows mean response values.

Displays a cube plot for factors in C...C. You can draw a cube plot with 2–8 factors for each response entered with the RESPONSE subcommand with data means or 2–7 factors with fitted means. The factor columns can be numeric or text, and can contain any values. The levels do not need to be in any special order. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on. All factors must have 2 levels, and can have center points.

When you use the RESPONSE subcommand, FFCUBE displays the means for the response at each point in the cube where observations are measured. When you do not use the RESPONSE subcommand, FFCUBE draws points on the cube for the effects that are in your model, which helps you visualize the design.

RESPONSE C...C
Plots the raw response data that is in columns C...C. The means for the cube points are calculated from the data. You can specify 2–8 factors.

If you use FMEANS, RESPONSE is required.

FMEANS C...C

Plots the least squares or the fitted means that are in columns C...C. You can specify 2-7 factors. The column must contain the fitted means in standard order for the cube points. The last row must be the overall mean of the data.

If you use FMEANS, RESPONSE is required.

INUNIT K (required)

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded form
1	Uncoded form

CTPT C

Specifies a center point column in C. Minitab considers a value of 0 to be a center point. A valid column contains only values of 0 and 1.

If you don't use MIXTURE, CTPT is required. You cannot use both MIXTURE and CTPT.

LEVELS K K, ..., K K (required)

Specifies the levels for the uncoded units for each factor in K K, ..., K K. The order of the parameters must be identical to the order of the factors that follow FFCUBE.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB

TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

MIXTURE
Specifies a mixture design, indicating that the factors are process variables.

Response Optimization for Mixtures

MIXOVER: Session command for creating an overlaid contour plot

MIXOVER
Draws an overlaid contour plot for multiple (up to 10) responses of mixture models.

Contour plots for mixture designs show how response variables are related to three continuous design variables, while holding the other variables in a model at certain settings. In a contour plot, the fitted response model is viewed as a 2-dimensional surface where all points that have the same fitted value are connected to produce contour lines of constants. Contour plots are useful for establishing operating conditions that produce desirable response values.

Applications that involve multiple responses present a different challenge than single-response experiments. Because each response is important in determining the quality of the product, you need to consider the responses simultaneously. Optimal settings for input variables for one response might be far from optimal for another response. Overlaid contour plots allow you to visually identify an area of compromise among the various responses.

For more information about graphics options for MIXOVER, go to Graphics options for MIXOVER on page 1233.

RESPONSE C...C
Specifies a set of up to 10 response variables for the experiment in C...C.

MODEL C...C
Specifies a set of model columns (one for each response), which contain design and model information in C...C. Each model column is stored by the MODEL subcommand of MIXREG.

SINGLE C C [C]
Specifies three components or two numerical process variables for a single plot. You must specify the components or process variables on the list in the model column.

LEVELS K...K

Specifies the values of contour levels in K...K. K must be between 2 and 15.

TOTAL K [K...K]

TOTAL C

Specifies single or multiple total values. A single K must equal the common row sum of components. Multiple Ks are accepted only if the Amount column is specified. The total subcommand must be specified if the Amount column's specified either with a single K = 1 or with multiple K values that match the values in the Amount column.

LOWER K...K

LOWER C

Specifies the lower bound of the components. The number of Ks must equal the number of components. All Ks must be 0 <= K < TOTAL and SUM(K) < TOTAL. (TOTAL is the first value on Total.)

UPPER K...K

UPPER C

Specifies the upper bound of the components in K...K. The number of Ks must equal the number of components. All Ks must be 0 < K <= TOTAL and SUM(K) > TOTAL. (TOTAL is the first value on Total.) Each upper bound must be greater than the corresponding lower bound.

CLINEAR K...K

CLINEAR C

Specifies the linear constraints on the components. The number of Ks must equal the number of components + 2. The first K, the lower bound, must be less than the last K, the upper bound. If one of the bounds is missing, set it to the missing value ' * '. You can use up to 10 CLINEAR subcommands.

PROPORTION

Specifies the proportion unit for the axes on the plot.

PSEUDO

Specifies the pseudo-component unit for the axes on the plot.

CLOW, CMIDDLE, and CHIGH

Sets the values of all the components that are not on the plot. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound. These commands are not available for the process variable contour plot.

Command	Bound or level
CLOW	Lowest bound
CMIDDLE	Middle level
CHIGH	Upper bound

CHOLD K...K

Holds components that are not on the plot to values that you specify in K...K. Each K must be between the lower and upper bounds (inclusive) of the corresponding component. Only design points that match these values are plotted. By default, all components that are not on the plot are held at their lower bound.

CMIN K...K

CMIN C

Specifies the minimum axis values for all components, including those that are not on the plot. The number of constants and the length of the column must match the number of components.

FLOW, FMIDDLE, and FHIGH

Sets the level of all the process variables that are not on plot. By default, all process variables are held at –1.

Command	Level
FLOW	Low level (–1)
FMIDDLE	Middle level (0)
FHIGH	High level (1)

FMIN K...K

FMIN C

Specifies the minimum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

FMAX K...K

FMAX C

Specifies the maximum axis values for all process variables, including those that are not on the plot. The number of constants or the length of the column must match the number of process variables.

AHOLD K

Sets the mixture amount at the specified value in K. Only design points that match these values are plotted. By default, the amount value is held at the first value in the Amount column.

SGRID [K]

Displays grid lines on each axis at specified intervals. All Ks must be 0 < K < 1. By default, K is 0.2.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MROPT: Session command for the Response Optimizer

MROPT C...C

Use the Response Optimizer to identify the combination of input variable settings that optimize a single response or that jointly optimize a set of responses. Joint optimization must satisfy the requirements for all the responses in the set. The overall desirability (D) is a measure of how well you satisfied the combined goals for all the responses. Overall desirability has a range of 0 to 1. A value of 1 represents the ideal case. A value of 0 indicates that one or more responses are outside their acceptable limits.

Minitab calculates an optimal solution and draws a plot. The optimal solution serves as the starting point for the plot. With the optimization plot, you can interactively change the input variable settings to perform sensitivity analyses and to possibly improve the initial solution.

Note Although numerical optimization, along with graphical analysis, can provide useful information, it is not a substitute for subject matter expertise. Be sure to use relevant background information, theoretical principles, and knowledge from observation or previous experimentation when you apply these methods.

Enter a set of response columns (up to 10) in the main command. You can use MIXTURE to match the responses with their appropriate column. This column contains the model information that MROPT needs to find the optimum solution and draw the trace plots.

For each response, you must do the following:

- Select a goal.
- Specify a minimum acceptable value, a maximum acceptable value, and a target value.
- Specify a utility transfer function value.
- Specify a weight or an importance level.

MODELS C...C

Specifies the model for the responses to be optimized. The arguments, one for each response, are the columns stored by the MODEL subcommand of MIXREG. If you use session commands to optimize responses, you must manually store the columns using MIXREG and MODEL.

LINEAR K...K

Specifies linear constraints for mixture components. You can use LINEAR up to 10 times.

For p mixture components, you must specify (p + 2) constant arguments. The first argument specifies the lower bound for a linear inequality constraint. The next p arguments are the coefficients of the linear function. The final argument is the upper bound on the linear function. Therefore, the subcommand actually specifies two linear inequality constraints. If one of the bound arguments is ' * ' then there is only one inequality, defined by the other bound. An inequality with an upper bound can be expressed in terms of a lower bound by multiplying all coefficients and the bound by –1.

Used only as a subsubcommand with MIXTURE.

TOTAL K

Specifies the total mixture amount or amounts. K is 1–5.

INUNIT K

Specifies whether the design is in coded form or uncoded form.

Value of K	Form of design
0	Coded
1	Uncoded

FLEVELS K K, ..., K K

Specifies values for the levels of each factor in K K,..., K K. The level values can be text or numeric.

GOAL K...K

GOAL C

Specifies a goal for each response. Enter the goal as a code from the following table in either a list of constants (K...K) or a column (C).

Code	Goal	Description
1	Smaller is better	Set the response as small as possible.
2	Target is best	Set the response as close to the target as possible.
3	Largest is best	Set the response as large as possible.

UTWEIGHTS K...K

UTWEIGHTS C

Specifies the weights for each response. The weight defines the shape of the desirability function for each response. For each response, you can select a weight (from 0.1 to 10) to emphasize or de-emphasize the target.

Weight	Emphasis
Less than 1 (minimum is 0.1)	Less emphasis the target
Equal to 1	Equal emphasis on the target and the bounds
Greater than 1 (maximum is 10)	More emphasis on the target

IMPORTANCE K...K

IMPORTANCE C

Specifies importance levels for each response. The levels are numbers between 0.1 and 10. If all responses are equally desirable, use the default value of 1.0 for each response. Enter one argument (K) for each response, or a column C) that contains the importance values.

The composite desirability is then the geometric mean of the individual desirabilities. However, if some responses are more important than others, then you can include this information in the optimal solution by setting unequal importance values.

Larger values correspond to more important responses. Smaller values correspond to less important responses. You can also change the weights to perform sensitivity analyses on the optimal solution. For example, the optimal solution when a response has a large weight might be very different from the optimal solution when the same response has a small weight.

START K...K

Specifies a starting point for the search algorithm by giving a coordinate for each factor. Enter the coordinates as constants. Each factor must have one value, and each value must be between the minimum and the maximum for that factor.

DSTORE C

Stores the composite desirability values that are calculated at the points in the original design in a column. You can use these values to produce wireframe plots or contour plots of the composite desirability surface using CONTOURPLOT.

OPTIPLOT

Displays a table that shows the numerically optimal solution, the desirabilities for the responses, and the interactive optimization plot.

Minimum, maximum, and target values

Specifies the minimum acceptable values, the maximum acceptable values, and the target values for the responses. You must use all three subcommands. All three commands are needed to obtain the numerically optimum solution. You must also enter a value for each response in each of these subcommands. In other words, for each response, you need a minimum acceptable value, a maximum acceptable value, and a target value. You can enter the values by using a list of constants, or a column.

If you have a smaller-is-better goal for a response, your goal is to minimize the response. Identify the point of diminishing returns. That is, you want to minimize the response, but going beyond a certain value makes a small difference or no difference. If such a point exists, use it as the target value. If it does not exist, use a very small number, one that is

probably not achievable, as the target value. For a smaller-is-better response, the target and the minimum acceptable value are also equal.

The same is true when the goal is larger-is-better. However, in the larger-is-better case, you need a value on the upper end of the range instead of on the lower end.

When the goal is target-is-best, you probably have upper and lower specification limits for the response. You can use these limits as minimum and maximum acceptable values. If not, you might want to use upper and lower points of diminishing returns, or some other method.

MINACCEPT K...K

MINACCEPT C
Specifies the minimum acceptable values.

MAXACCEPT K...K

MAXACCEPT C
Specifies the maximum acceptable values.

TARGET K...K

TARGET C
Specifies the target values.

Taguchi Designs

OADESIGN: Session command for creating a Taguchi orthogonal array design

OADESIGN K K K [K K]
Creates a Taguchi orthogonal array design, where the first K specifies the number of runs in the design and the remaining Ks specify number of levels and number of factors, respectively.

Use OADESIGN to specify a Taguchi orthogonal array design. OADESIGN can take either 3 or 5 integer arguments, which correspond to the standard Taguchi notation. For example, a commonly used orthogonal array is denoted by L18 (2**1 3**7). This means that the design has 18 runs, with one two-level factor, and seven three-level factors. The Minitab command that will specify this design is as follows:

```
OADESIGN 18 2 1 3 7.
```

Or, as follows:

```
OADESIGN 18 3 7 2 1.
```

If you want to use this array with one two-level factor and six three-level factors, the command is as follows:

```
OADESIGN 18 2 1 3 6.
```

To store your design for additional analysis, use the DESIGN subcommand. (The DESIGN subcommand replaces the obsolete subcommands XMATRIX and SIGNAL.)

LEVELS K K, ..., K K
Specifies the levels for each factor in K K, K K ... K K.

If you do not specify levels, Minitab uses the default levels given in the catalogue of arrays. For example, the default levels of a 3-level column are 1, 2, and 3.

The following example shows how LEVELS is used:

```
OADESIGN 18 3 4 2 1;
DESIGN;
LEVELS 1 2 3 "red" "green" "blue" 2.7 2.9 3.4 3.9 4.2 6.8 -1 1.
```

This design will use five columns of the L18 (2**1 3**7) array. The first four factors, which will be stored in columns C1-C4, will have 3 levels each. The fifth factor, which will be stored in column C5, will have 2 levels. The levels of factor C1 will be 1, 2, and 3. The levels of C2 will be "red", "green", and "blue". The levels of factor C3 are 2.7, 2.9, ad 3.4.

Note It is acceptable to repeat levels for the same factor, as long as there are at least two distinct levels. This allows the use of "dummy treatments", an important technique for customizing Taguchi designs. For example, the 8-level column in the L16 (8**1 2**8) array can be used for a 7-level factor by specifying levels 1 1 2 3 4 5 6 7.

COLUMNS K...K

Specifies which columns of the full orthogonal array are used in the design in K...K.

Consider the following example:

```
OADESIGN 18 3 4 2 1;
DESIGN;
COLUMNS 2 3 5 7 1.
```

Worksheet column C1 will be the same as column 2 in the L18 (2**1 3**7) array. Column C2 will be the same as column 3; C3 will be column 5; C4 will be column 7; and C5 will be column 1.

INTERACTIONS C C, ..., C C

Specifies which interactions among factors in C C you want to estimate. INTERACTIONS requests that Minitab automatically assign factors to columns in such a way that the 2-way interactions listed are not confounded with any of the main effects or with each other. If you want to remove an interaction from your model, you should run the OADESIGN command to re-specify your model with the appropriate terms.

DESIGN

Names and stores the standard order, run order, point type, block assignment, and each factor in separate columns. Unless you have only 1 empty worksheet in the project, Minitab stores the columns in a new worksheet.

NAMES K K, ..., K K

Specifies names for the factors. Enclose names in quotation marks. If you do not use NAMES, then the names of the factors are letters, skipping the letter I: A, B, C...

SNAME K

Use SNAME to add a signal factor to create a dynamic Tauchi design. Enter the name within double quotation marks. When you specify SNAME, also specify SLEVEL.

SLEVEL K...K

Use SLEVEL specify the levels of the signal factor for a dynamic Taguchi design. Levels for a signal factor must be numeric. Minitab replicates the orthogonal array part of the design at each level of the signal factor.

Consider the following example:

```
OADesign 4 2 3;
Design;
SNAME "Signal";
SLEVEL 1:5/2.
```

The design stored in the worksheet would be as follows:

C1	C2	C3	C4
A	B	C	Signal
1	1	1	1
1	1	1	3
1	1	1	5
1	2	2	1
1	2	2	3
1	2	2	5
2	1	2	1
2	1	2	3
2	1	2	5
2	2	1	1
2	2	1	3
2	2	1	5

In general, the total number of runs will be the number of rows in the orthogonal array times the number of levels in the signal variable.

ROBUST: Session command for analyzing a Taguchi design

ROBUST C = C...C

Analyzes a Taguchi design with responses in C and factors and interactions in C...C.

Calculates signal-to-noise ratios for an inner-outer array design and generates main effects and interactions plots of the signal-to-noise ratios, means, slopes, or standard deviations versus the control factors in the inner array. Use response tables and ANOVA results to determine what factors and interactions are important and which factor settings minimize variability.

To use **Predict Taguchi Results** on the Minitab menu, use DESIGN and DOE to store the model information.

DESIGN C...C

Specifies all of the factors in your design in C...C if you have omitted some factors in the model specified in ROBUST.

DOE

Stores model information to enable **Predict Taguchi Results** on the Taguchi menu. You can make predictions with OAPREDICT whether or not you store the model information.

DYNAMIC C

Specifies that the signal factor is in C. DYNAMIC cannot be used with any subcommands that apply only to static response analysis: NOMONE, NOMTWO, LOWER, HIGHER, GMEANS, MEANS, CV, TMEANS.

Graphs

GSN

Displays main effects for s/n ratios. GSN plots the S/N ratio as the response variable. Choose the appropriate S/N ratio with the NOMONE, NOMTWO, LOWER, or HIGHER subcommands (the default is NOMTWO).

GMEANS

Displays main effects for means. GMEANS cannot be used with DYNAMIC. You can use as many GMEANS as you want. See LOGSTD if you want to use the log-transformed standard deviations as the response variable.

GSTDS

Displays main effects for standard deviations. You can use as many GSTDS as you want. See LOGSTD if you want to use the log-transformed standard deviations as the response variable.

GSLOPES

Displays main effects for slopes. GSLOPES can only be used with DYNAMIC. You can use as many GSLOPES as you want. See LOGSTD if you want to use the log-transformed standard deviations as the response variable.

GSCATTER

Displays scatterplots of data with fitted lines. Use GSCATTER to request scatterplots of the responses versus the signal. GSCATTER can be used only with DYNAMIC.

INTROWS C...C

Specifies factors to be used as the grouping variable in the interaction plot matrix. The plots in the same row share the factor that is used as the grouping variable.

INTCOLS C...C

Specifies factors to use for the horizontal scale of the interaction plot matrix. The plots in the same column share the factor used for the horizontal scale.

INTSEP

Generates separate interaction plots for interactions specified in the model.

RTYPE K

Specifies the type of residual to plot with the graph subcommands.

Value of K	Type of residuals
1 (default)	Regular or raw residuals (RESIDUALS)
2	Standardized residuals (SRESIDUALS)
3	Deleted Studentized residuals (TRESIDUALS)

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER [C]

Plots the residuals versus the order of the data or vs the order specified in [C]. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK [C]

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data or vs order specified in [C].

GVARIABLE C...C

Plots the residuals versus variables. Displays a separate graph for the residuals versus each specified column.

Analysis

TMEANS

Generates a response table for means. TMEANS cannot be used with DYNAMIC.

TSN

Generates a response table for s/n ratios.

TSTDS

Generates a response table for standard deviations.

TSLOPES

Generates a response table for slopes. TSLOPES can only be used with DYNAMIC.

AMEANS

Generates ANOVA results for means. AMEANS cannot be used with DYNAMIC.

ASN

Generates ANOVA results for s/n ratios.

ASTDS

Generates ANOVA results for standard deviations.

ASLOPES

Generates ANOVA results for slopes ratios. ASLOPES can only be used with DYNAMIC

Options

NOMONE

Specifies the s/n ratio based on standard deviations only.

Use NOMONE when the goal is to target the response. The formula is $S / N = -10 (\log_{10}(s^2))$.

NOMTWO

Specifies the s/n ratio based on means and standard deviations (default).

Use NOMTWO when the goal is to target the response. The formula is $S / N = 10 (\log_{10} (Xbar^2 / s^2))$.

LOWER

Specifies the s/n ratio when the goal is to minimize the response.

Use LOWER when the goal is to minimize the response. The formula is $S / N = 10 (\log_{10} (\Sigma Y^2 / n))$.

HIGHER

Specifies the s/n ratio when the goal is to maximize the response.

Use HIGHER when the goal is to maximize the response. The formula is

$S / N = -10 (\log_{10}(\Sigma (1 / Y^2) / n))$

LOGSTD

Specifies the loge transformation for standard deviations. If you want to produce main effects and interactions plots using the standard deviation as the response variable, consider using a loge transformation on the standard deviations to stabilize their variability.

YREF K

Specifies the response reference value for dynamic designs.

Specifies the reference point through which the regression line must pass in a dynamic response model. If the subcommand is not used, the reference value is taken to be zero by default. If neither YREF nor SREF are used, the regression line passes through the origin.

YREF can only be used with DYNAMIC.

SREF K

Specifies the signal reference value for dynamic designs.

Specifies the reference point through which the regression line must pass in a dynamic response model. If the subcommand is not used, the reference value is taken to be zero by default. If neither YREF nor SREF are used, the regression line passes through the origin.

SREF can only be used with DYNAMIC.

NOREF

Specifies no reference values for the regression. Minitab estimates a separate regression intercept for each run.

Storage

You must use storage commands with one or more of the commands AMEANS, ASLOPES, ASNS, or ASTDS.

SNRATIO C

Stores s/n ratios in C.

MEANS C

Stores means in C. MEANS cannot be used with DYNAMIC.

SLOPES

Stores slopes in C. SLOPES can only be used with DYNAMIC.

INTERCEPTS

Stores intercepts in C. INTERCEPTS can only be used with DYNAMIC.

STDEVS

Stores standard deviations in C.

LSTDEVS

Stores natural log of standard deviations in C.

CV
Stores coefficients of variation in C. CV cannot be used with DYNAMIC.

FITS C
Stores the fitted values.

RESIDUALS C
Stores the residuals (fitted values – observed values).

SRESIDUALS C
Stores the standardized residuals.

TRESIDUALS C
Stores the deleted Studentized residuals in C.

COEFFICIENTS C
Stores the estimated coefficients.

XMATRIX *M*
Stores the design matrix for the regression model in M.

HI C
Stores the leverage values.

COOK C
Stores Cook's distance.

DFITS C
Stores the DFITS (also called DFFITS) values.

OAPREDICT: Session command for calculating predicted response values

OAPREDICT C...C = C...C
Predicts response values for statistics stored in C...C for factors and interactions in C...C.

Calculate predicted response values after you have analyzed a Taguchi design and examined the response tables, ANOVA results, and plots to determine which factor settings should achieve a robust product design. OAPREDICT allows you to predict S/N ratios and response characteristics for selected factor settings.

To use OAPREDICT, first store response characteristics (S/N ratios, means, slopes, standard deviations) you want to predict with ROBUST. Specify stored response characteristic columns on left-side of model. Specify factors and interactions on the right-side of the model.

For example, you might choose the best settings for the factors that have the greatest effect on the S/N, and then wish to predict the S/N and mean response for several combinations of other factors. OAPREDICT provides the expected responses for those settings. You should choose the results that come closest to the desired mean without significantly reducing the S/N ratio. You should then perform a follow-up experiment using the selected levels, to determine how well the prediction matches the observed result.

If there are minimal interactions among the factors or if the interactions have been correctly accounted for by the predictions, the observed results should be close to the prediction, and you will have succeeded in producing a robust product. On the other hand, if there is substantial disagreement between the prediction and the observed

results, then there may be unaccounted for interactions or unforeseen noise effects, which may require further investigation.

LABELS K...K
Specifies labels for the predicted values in the output. The arguments must be text constants.

LEVELS *E...E*
Specifies the factor levels for which you want to calculate predicted response values. The arguments may be either constants or columns. The levels must agree with the levels in the factor columns.

PFITS C...C
Stores the predicted values in specified columns. The number of arguments should match the number of response characteristics listed on the left-hand side of the model in the OAPREDICT command.

UNITS K K
Specifies the units of the inner array factors and the units for the new levels, 0 for coded and 1 for uncoded. The first argument indicates the units of the inner array columns and the second argument indicates the units of the levels given by the LEVELS subcommand.

NLEVELS K...K
Specifies the number of levels for each factor. Use when the two arguments of UNITS are different. It is used in conjunction with FLEVELS to match coded levels with uncoded levels for each factor.

FLEVELS K...K
Specifies the uncoded levels of the factors listed in the model. Use FLEVELS if the two arguments of UNITS are different. The number of arguments for FLEVELS should be the sum of the arguments of NLEVELS. NLEVELS is used to determine which levels correspond to which factors.

Example:

```
OAPREDICT 'SNRA1' = C1 C3-C5;
  LEVELS 2 1 1 1;
  UNITS 1 0;
  NLEVELS 2 3 3 3;
  FLEVELS "low" "high" "low" "low" "high" 1 2 3 1 2 3.
```

In this example, OAPREDICT would predict a new value of the s/n ratio by fitting a main effects model based on the four inner array factors C1, C3–C5. The UNITS subcommand indicates that the inner array is displayed in uncoded units, and the new levels are specified in coded units. The NLEVELS subcommand indicates that C1 has 2 levels and C3–C5 have 3 levels each. The FLEVELS subcommand indicates that for factor C1, the uncoded units are "low" and "high"; for factor C3, the uncoded units are "low", "low" and "high", i.e. coded levels 1 and 2 would both correspond to uncoded level "low", etc.

Modify and Display Designs

MDESIGN: Session command to modify the design properties or display of a design in the worksheet

MDESIGN
The modifications that are available depend on the type of design in the worksheet. You can use NEWWS to put the modified design in a new worksheet for all types of designs. Generally, you can make only 1 modification to

the design each time you use MDESIGN. For detailed explanations, including instances where you can make multiple modifications, go to the section on the subcommands.

Modifications for screening designs

Modification	Subcommand
Modify factor names or levels	VAR
Replicate the design	REPLICATE
Randomize the design	RANDOMIZE
Renumber the design	RENUMBER
Display the points in run order or standard order	SORT
Display the levels in coded or uncoded units	CODED and UNCODED
Fold the design (Plackett-Burman only)	FOLD
Add axial points (Plackett-Burman only)	AXIAL

Modifications for factorial designs

Modification	Subcommand
Modify factor names or levels	VAR
Replicate the design	REPLICATE
Randomize the design	RANDOMIZE
Renumber the design	RENUMBER
Display the points in run order or standard order	SORT
Display the levels in coded or uncoded units	CODED and UNCODED
Fold the design (2-level fractional factorial designs and Plackett-Burman designs)	FOLD
Add axial points (2-level fractional factorial designs and Plackett-Burman designs)	AXIAL

Modifications for response surface designs

Modification	Subcommand
Modify factor names or levels	VAR
Replicate the design	REPLICATE
Randomize the design	RANDOMIZE
Renumber the design	RENUMBER
Display the points in run order or standard order	SORT
Display the levels in coded or uncoded units	CODED and UNCODED

Modifications for mixture designs

Modification	Subcommand
Rename mixture components, rename process variables, or change the levels of process variables	VAR
Replicate the design	REPLICATE
Randomize the design	RANDOMIZE
Renumber the design	RENUMBER
Display the points in run order or standard order	SORT
Display the levels of the process variables in coded or uncoded units	CODED and UNCODED
Display the levels of the mixture components in amount, proportion, or pseudocomponent units	AMOUNT, PROPORTION, and PSEUDO

Modifications for Taguchi designs

Modification	Subcommand
Modify factor names or levels	VAR
Display the levels in coded or uncoded units	CODED and UNCODED
Add a signal factor to a static design, ignore an existing signal factor (treat the design as static), or add new levels to an existing signal factor.	SIGNAL

Subcommands

NEWWS

Use NEWWS to put the modified design in a new worksheet. Without NEWWS, the modification overwrites the columns in the existing worksheet. You can use NEWWS in combination with any of the other commands.

VAR C

Use VAR to change the name or levels of a factor, process variable, or mixture component. Use C to specify which column to modify. You can specify VAR multiple times for a single instance of MDESIGN. You can use NAME and LEVELS at the same time. You cannot use LEVELS for mixture components.

NAME K

Use NAME to specify the new name for a factor, process variable, or mixture component. Enter the name within double quotation marks.

LEVELS K ... K

Specify the new uncoded levels for the factor or process variable that you specify with VAR. For a factor with low and high levels, the first K is the low level and the second K is the high level. Enter text levels within double quotation marks. For a text factor with more than 2 levels, as in a general full factorial design, specify the new levels in the same order as in the command language for the creation of the design.

REPLICATE K

Use REPLICATE to reproduce the rows in the current worksheet K times. If the original design is all in 1 block, each replicate is a new block. If the original design has multiple blocks, each replicate contains the same number of blocks as the original design.

Because the order from the worksheet is the same for each replicate, modify the design a second time to randomize it after you add the replicates. Use the RANDOMIZE subcommand with a new instance of MDESIGN. For more information, go to the RANDOMIZE subcommand.

RANDOMIZE

Use RANDOMIZE to replace the run order column with a new, random order. You can use RANDOMIZE and SORT at the same time to put the design in the new run order.

BASE K

Use BASE to get the same random order when you modify the design again. Enter a positive integer as the seed for the random number generator.

WP K

Use WP with split-plot designs. The randomization affects only the rows where the value in the whole plot column equals K.

BLOCKS K

Use BLOCKS with a blocked design. The randomization affects only the rows where the value in the block column equals K. For example, if you add a replicate to a design after you collected some response data, you can use BLOCKS to randomize only the rows in the block for the new replicate.

RENUMBER

Use RENUMBER to make the run order the order of the rows in the worksheet.

SORT C

Use SORT to sort the worksheet by the column that you specify. You can specify only the column that contains the standard order for the design or the column that contains the run order for the design. By default, the standard order is in column C1 and the run order is in column C2. Minitab adds missing value symbols to columns that are shorter than the design columns so that the rows can be in the new order. SORT does not affect columns that are longer than the design.

EXCLUDE C...C

Use EXCLUDE to preserve the order of columns that you do not want to sort.

CODED

When the worksheet shows uncoded units, specify CODED to change to coded units.

UNCODED

When the worksheet shows coded units, specify UNCODED to change to uncoded units.

FOLD [C]

Use FOLD with 2-level fractional factorial designs or Plackett-Burman designs. FOLD folds the design on the specified factor. To fold on multiple factors, use FOLD once for each factor within the same instance of MDESIGN. If you do not include a factor, Minitab folds the design on all factors.

The signs on the factors that you specify are switched. The signs on the remaining factors are not switched. These rows for the non-specified factors are then appended to the end of the design matrix, which doubles the number of runs.

Consider the following example.

```
FFDESIGN 3 4;
DESIGN.
MDESIGN;
   FOLD A.
```

First, Minitab generates a 3 factor design in 4 runs. The first 4 rows of the data matrix below are these runs. Then, Minitab folds on factor A. This adds 4 additional runs.

A	B	C
–	–	+

+	–	–
–	+	–
+	+	+
+	–	+
–	–	–
+	+	–
–	+	+

Folding is a way to reduce confounding. If you fold on one factor, say A, then A and all its two-factor interactions are free from other main effects and two-factor interactions. If you fold on all factors, then all main effects are free from all two-factor interactions.

When you fold, the defining relation is usually shortened. Specifically, any word in the defining relation that has an odd number of the letters on which you folded the design is omitted. If you fold a design and the defining relation is not shortened, then the folding adds replicates, but it does not reduce confounding. In this case, Minitab displays an error message.

If you fold a design that is blocked, the same block generators are used for the folded design as for the unfolded design.

AXIAL

Use AXIAL with 2-level factorial designs and Plackett-Burman designs. To use AXIAL, the design must have only continuous factors. Use AXIAL to change the design into a central composite response surface design by adding a block of axial points.

ALPHA K

Use ALPHA to specify the alpha for the central composite design. Alpha is the distance from the center of the design to the axial points in coded units. If you do not use ALPHA, Minitab chooses alpha so that the design is as rotatable as possible with orthogonal blocks. Alpha must be positive. A value of 1 results in a face-centered design. A value that is less than 1 places the axial points inside the cube. A value that is greater than 1 places the axial points outside the cube.
Central composite designs on page 1219 shows the values of alpha that Minitab uses when Minitab creates a central composite design.

FACE

Use FACE to specify an alpha value of 1. This alpha creates a face-centered central composite design. Alpha is the distance from the center of the design to the axial points in coded units.

CENTER K

Use CENTER to specify the number of center points to add to the axial block.
Central composite designs on page 1219 shows the number of center points that Minitab includes in the axial block when Minitab creates a central composite design.

AMOUNT

When the worksheet shows proportion or pseudocomponent units for a mixture design, use AMOUNT to change to amount units.

PROPORTION

When the worksheet shows amount or pseudocomponent units for a mixture design, use PROPORTION to change to proportion units.

PSEUDO

When the worksheet shows proportion or amount units for a mixture design, use PSEUDO to change to pseudocomponent units.

SIGNAL

Use SIGNAL to add, add levels to, or ignore a signal factor in a Taguchi design. Use the sub-subcommands to specify the use.

INCLUDE

Use INCLUDE to add a signal column to a Taguchi design. Also use SNAME and SLEVEL to specify the name and levels of the signal factor. Omit INCLUDE when you add levels to a signal factor.

SNAME K

Use SNAME to name the signal factor that you add to a design. Enter the name within double quotation marks.

SLEVEL K...K

Use SLEVEL with INCLUDE to specify the levels for a signal factor that you add to a design. Use SLEVEL without INCLUDE to add levels to an existing signal factor. Levels for a signal factor must be numeric. Minitab replicates the orthogonal array part of the design at each level of the signal factor.

IGNORE

Use IGNORE to change a dynamic Taguchi design into a static Taguchi design. The column for the signal factor that is in the worksheet does not change.

Quality and Process Improvement

Quality Planning Tools

FISHBONE: Session command for creating a cause-and-effect diagram

FISHBONE

Generates a cause-and-effect diagram (also called a fishbone diagram).

Branches

Specifies lists of causes for the corresponding branches and subbranches of the diagram. You can use these commands repeatedly and specify up to 10 branches.

When data are in columns, each column contains the names of the causes for a branch of the diagram. Entries in the columns can be up to 72 characters wide. The first 20 characters of each entry in the column display on the diagram. Odd branches are on top (5, 3, then 1, left to right), and even branches on the bottom (6, 4, then 2).

BRANCH K C

Use BRANCH K C and SUBBRANCH K C when the causes are listed in columns.

SUBBRANCH K C

Use BRANCH K C and SUBBRANCH K C when the causes are listed in columns.

BRANCH K K

Use BRANCH K K and SUBBRANCH K K when you type the causes as text.

SUBBRANCH K K

Use BRANCH K K and SUBBRANCH K K when you type the causes as text.

Other subcommands

NAME K *"text"*

By default, the labels for branches 1 through 6 are (in order) Personnel, Machines, Materials, Methods, Measures, and Environment. NAME K "text" changes the default labels for K branch. Odd branches are on top (5, 3, then 1, left to right), and even branches on the bottom (6, 4, then 2).

NOEMPTY C

Suppresses branches that you have not specified data for with BRANCH 1, BRANCH 2, BRANCH 3, BRANCH 4, BRANCH 5, and BRANCH 6. When you do not specify NOEMPTY, empty branches still display.

NOLABELS

Suppresses the default branch labels. Use NOLABELS to draw a blank diagram.

EFFECT *"text"*

Displays the name of the effect, or problem, that is being solved. The text string can be up to 72 characters in length. The first 20 (approximately) characters display on the diagram.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB

TIF grayscale

TIF

TIF color

BMPB

BMP grayscale

BMPC

BMP color

GIF

GIF

EMF

EMF

RESOLUTION K

Saves the graph at a resolution of K dots per inch.

MVARCHART: Session command for creating a multi-vari chart

MVARCHART

Creates multi-vari charts for up to four factors. Multi-vari charts are a way of presenting analysis of variance data in a graphical form to provide a visual alternative to analysis of variance. You can use a multi-vari chart in the preliminary stages of data analysis to see the data. The chart displays the means at each factor level for every factor.

RESPONSE C

Specifies the numeric response column C.

FACTORS C...C

Specifies one to four factor columns.

DATAON

Displays individual data points on the plot.

MUAON

Displays connect lines for the means for factor A.

MUBON

Displays connect lines for the means for factor B.

MUCON

Displays connect lines for the means for factor C.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

NMVARCHART: Session command for creating variability charts

NMVARCHART

Creates variability charts for up to 8 factors. Use these charts to visualize different sources of variation. You can use variability charts in the preliminary stages of data analysis to see the data. For example, you can view analysis of variance data or variation from a measurement system graphically.

The command can produce two charts. The first chart, the average response with variation chart, is shown by default. This chart plots the original data points and the range of each cell formed by the combination of levels from all factors. This chart also optionally displays the means for each factor level. Use to evaluate measurement consistency within cells and to see how different cell means are from each other.

The second chart, the standard deviation chart, plots the standard deviations for each cell formed by the combination of levels from all factors.

RESPONSE C

Specifies the numeric response column C.

FACTORS C...C

Specifies one to eight factor columns. The first C is the first factor, the second C is the second factor, and so on.

The factor columns must be the same length as the response column.

DATAON

Displays individual data points of each cell, defined by the factors, on the average chart.

RANGE
Displays every range bar that connects the min and max of the data points within each cell.

CELLON
Connects the cell means.

MUAON
Displays level means for factor A.

MUBON
Displays level means for factor B.

MUCON
Displays level means for factor C.

MUDON
Displays level means for factor D.

MUEON
Displays level means for factor E.

MUFON
Displays level means for factor F.

MUGON
Displays level means for factor G.

OMEAN
Displays the overall mean.

GSTD
Displays the standard deviation chart.

SMEAN
Displays the mean of the standard deviations.

PARETO: Session command for creating a Pareto chart

PARETO C
Generates a Pareto chart for one sample, or one for each group when you use a BY variable.

You can use PARETO in either of two ways. First, you enter a column of raw data, where each observation is an occurrence of a type of defect. PARETO counts the number of occurrences for each unique defect, then uses the counts to construct the chart. Second, you enter a list of defect names in a column as the PARETO argument and a list of counts in a second column as the COUNTS argument. The defect name and count in the same row are paired in the plot. In both cases, the defects can be either text or numeric data. Thus, you can use either the actual defect names, or you can use numeric codes that represent the defects.

OTHERS K
By default, PARETO generates bars for defects until the cumulative percent surpasses 95. PARETO groups remaining defects into a bar named "Others." Use OTHERS to specify a value other than 95 for the cumulative percent. For example, OTHERS 90 plots all defects until the cumulative percent surpasses 90, then lumps the rest into "Others."

NOOTHERS

Use NOOTHERS if you do not want to combine defects into an "Others" category.

COUNTS C

Instead of entering raw data, you may already know how many of each defect there are. In that case, enter the defect names in one column and the counts in another. Give the column of names in the main command line and the column of counts with COUNTS. PARETO then matches up the name and count in the same row when drawing the chart. The data in C must be integers.

BY C

Use BY to generate a separate Pareto chart for each unique value in the specified BY column. All graphs display together and use the same y-axis scale. When you do not use SEPARATE or INDEPEND, bars in all graphs display in the same x-axis order when you use BY. You can specify up to 25 by variable levels.

SEPARATE

When you use BY with SEPARATE, each graph displays in a separate graph, one after another, in the output pane. All graphs use the same y-axis scale and x-axis order.

INDEPEND

When you use BY with INDEPEND, the graphs are independent. Although all graphs use the same y-axis scale, each graph has an independent x-axis order. Thus, the bars in each graph display in Pareto order—from largest to smallest. When you use BY with INDEPEND, each graph displays in a separate graph, one after another, in the output pane.

YLABEL *"text"*

Specifies the y-axis default label of "Count".

XLABEL *"text"*

Specifies the x-axis default label of "Defect".

NOCUMULATIVE

Suppresses the cumulative symbols and line and percent scale.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

RUNCHART: Session command for creating a run chart

RUNCHART C C

RUNCHART K
Draws a run chart with tests for randomness based on runs.

RUNCHART shows the spread of your process data, and performs two tests for non-random behavior.

The command is entered differently depending on how the data is structured in the worksheet. For details, go to How to enter subgroup data on page 1246. RUNCHART can be used to plot subgroup means, subgroup medians, or the individual observations.

Regardless of subgroup size, RUNCHART plots all the individual observations versus the subgroup number. When you use a subgroup size greater than 1, RUNCHART also plots the subgroup data versus the subgroup number. Minitab connects the subgroup data with a line. The subgroup mean is plotted by default. Use MEDIANS to plot the subgroup medians instead of the subgroup means. A horizontal reference line shows the median of all observations.

RSUB C...C [K]
Enters subgroup data in rows. Enter a series of columns. You can optionally specify the number of models you would like. For example, if you have data across rows C1–C6 and you have two models (C1–C3 and C4–C6), you can enter `RSUB C1-C6 2` to produce two run charts. The number of models you specify must

divide evenly into the number of columns on the RSUB subcommand. For details, go to How to enter subgroup data on page 1246.

MEDIANS
Plots the subgroup medians rather than the subgroup means as points on the graph. If you do not use this subcommand, Minitab uses the subgroup means as the points on the graph. The points you choose to plot are used to count the number of runs for the tests for randomness.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

SYMPLOT: Session command for creating a symmetry plot

SYMPLOT C...C
Symmetry plots can be used to assess whether sample data come from a symmetric distribution. Many statistical procedures assume that data come from a normal distribution. However, many procedures are robust to violations of normality, so having data from a symmetric distribution is often sufficient. Other procedures, such as nonparametric methods, assume symmetric distributions rather than normal distributions. Therefore, a symmetry plot is a useful tool in many circumstances.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

Control Charts

Box-Cox Transformation

BOXCOX: Session command for performing a Box-Cox transformation

BOXCOX [C...C *E*]
Performs a box-cox transformation when the data are in C...C and the subgroups = E.

The Box-Cox transformation can be useful for correcting both nonnormality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, you do not need to correct for nonnormality unless the data are highly skewed. Some authors suggest that you should not transform data that are used in control charts, because control charts work well in situations where data are not normally distributed. They give an excellent demonstration of the performance of control charts when data are collected from a variety of non-symmetric distributions.

You can enter data when data are in columns or when data are in subgroups across rows. For details, go to How to enter subgroup data on page 1246.

You can use BOXCOX only with positive data.

BOXCOX can be used several ways:

- To estimate the best lambda value for the transformation, use BOXCOX with no subcommand.
- To estimate the best lambda value for the transformation, transform the data, and store the transformed data in the column or columns that you specify, use BOXCOX with STORE.
- To transform the data with a lambda value that you enter, and store the transformed data in a columns you specify, use BOXCOX with LAMBDA and STORE.

RSUB C...C
Specifies when data are in subgroups across rows.

LAMBDA K...K

LAMBDA C...C

Specifies your own value of lambda to use in the transformation. When you use LAMBDA, Minitab bypasses the lambda estimation step and transforms the data using the lambda value you specify. You must use the STORE subcommand with LAMBDA to see the transformed data in a column.

The within standard deviation is used in the estimation of lambda when you use optimal lambda.

STORE C...C

Stores the transformed data in column C...C.

Variables Charts for Subgroups

XRCHART: Session command for creating an Xbar-R chart

XRCHART [C...C *E*]

Creates an Xbar and R chart for subgroup means and ranges with subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The mean of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation is estimated by the average of subgroup ranges. You can optionally enter an historical standard deviation with the SIGMA subcommand. Base your estimate on the pooled standard deviation with the POOLED subcommand.

You can run the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

RSUB C...C

Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

BIASED

Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOLOWER K

Sets lower control limit bound for Xbar chart.

LOUPPER K

Sets upper control limit bound for Xbar chart.

VALOWER K

Sets lower control limit bound for R chart.

VAUPPER K

Sets upper control limit bound for R chart.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Note For Xbar charts and I charts, with extended control limits, only Test 1 is available. Test 1 fails when a point is outside the extended control limits.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

Note RCHART supports only tests 1 through 4.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Extended Limits

EXTENDEDLIMITS

EXTENDEDLIMITS calculates control limits based on an extended limits method. If you do not specify a subcommand for the calculation method, then Minitab uses `BWITHIN`. You can use either `BWITHIN` or `MINMAX`, but not both.

BWITHIN [K1][K2]

Use the between/within method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the false alarm probability for the between-subgroup standard deviation (alpha-between). Specify a numeric value between 0 and 1. The default value is approximately `0.1336144`.

MINMAX [K1][K2]

Use the min/max method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the number of smallest and largest subgroup means to use to estimate the minimum and maximum process locations. Specify an integer from 1 through the number of subgroups. The default is `3`.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C C
Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

XSCHART: Session command for creating an Xbar-S chart

XSCHART [C...C *E*]
Creates an Xbar and S chart for subgroup means and subgroup standard deviations with subgroup size = E.

If the subgroups are in columns, enter data with the CSUB subcommand. If the subgroups are in rows, enter data with the RSUB subcommand. For details, go to How to enter subgroup data on page 1246. The X chart is drawn in the upper half of the screen; the S chart in the lower half. Seeing both charts together allows you to examine both process level and process variation at the same time.

The process standard deviation is estimated the average of the subgroup standard deviations. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand, or base your estimate on the pooled standard deviations with the POOLED subcommand.

You can run the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, the column can contain different values for each historical group. If you enter one constant, all historical stages use that value. Otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, the column can contain different values for each historical group. If you enter one constant, all historical stages use that value. Otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOLOWER K

Sets lower control limit bound for Xbar chart.

LOUPPER K

Sets upper control limit bound for Xbar chart.

VALOWER K

Sets lower control limit bound for S chart.

VAUPPER K

Sets upper control limit bound for S chart.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes

do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Note For Xbar charts and I charts, with extended control limits, only Test 1 is available. Test 1 fails when a point is outside the extended control limits.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

Note SCHART supports only tests 1 through 4.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal l for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, Minitab finds the optimal value of lambda.

Extended Limits

EXTENDEDLIMITS

EXTENDEDLIMITS calculates control limits based on an extended limits method. If you do not specify a subcommand for the calculation method, then Minitab uses `BWITHIN`. You can use either `BWITHIN` or `MINMAX`, but not both.

BWITHIN [K1][K2]

Use the between/within method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the false alarm probability for the between-subgroup standard deviation (alpha-between). Specify a numeric value between 0 and 1. The default value is approximately `0.1336144`.

MINMAX [K1][K2]

Use the min/max method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the number of smallest and largest subgroup means to use to estimate the minimum and maximum process locations. Specify an integer from 1 through the number of subgroups. The default is `3`.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C C
Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

BWCHART: Session command for creating an I-MR-R/S chart

BWCHART [C...C *E*]
Creates a three-way control chart using both between-subgroup and within-subgroup variations, with subgroup size = E. You can enter data arranged in a single column or across rows. For more information, go to How to enter subgroup data on page 1246. When data are arranged across rows, you can optionally specify the number of models you would like. For example, if you have data across rows C1–C6 and you have two models (C1–C3 and C4–C6), you can enter `RSUB C1-C6 2` to produce two control charts. The number of models you specify must divide evenly into the number of columns on the RSUB subcommand.

In BWCHART, setup the data in several columns, with each row representing a subgroup. For example, if you took five measurements of each part, but in different locations, each column would contain the measurements from one of the locations, and each row would contain the measurements from a single part.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Use MU to estimate a value for the historical mean. The value that you enter can be either an actual known process parameter or an estimate that was obtained from past data.

If you use a column, it can contain different values for each historical stage. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical stage.

If you do not specify a value, Minitab estimates it from the data.

BSIG K...K

BSIG C

Use BSIG to estimate a value for the historical between-subgroup standard deviation. The value that you enter can be either an actual known process parameter or an estimate that was obtained from past data.

If you use a column, it can contain different values for each historical stage. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical stage.

If you do not specify a value, Minitab estimates it from the data.

WSIG K...K

WSIG C

Use WSIG to estimate values for the historical within-subgroup standard deviation. The value that you enter can be either an actual known process parameter or an estimate that was obtained from past data.

If you use a column, it can contain different values for each historical stage. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical stage.

If you do not specify a value, Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using EXCLUDE 1), then decide to omit observation 2 (using OMIT 2). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

RBAR

RBAR, SSBAR, and SPOOLED are mutually exclusive subcommands that control the estimation of the within-subgroup standard deviation when 1 < subgroup size <= 8.

RBAR uses the average of the subgroup ranges to estimate standard deviation. This method is the one that appears in many books and manuals that describe control charts.

SSBAR

RBAR, SSBAR, and SPOOLED are mutually exclusive subcommands that control the estimation of the within-subgroup standard deviation when 1 < subgroup size <= 8.

SSBAR uses the average of the subgroup standard deviations to estimate standard deviation.

SPOOLED

RBAR, SSBAR, and SPOOLED are mutually exclusive subcommands that control the estimation of the within-subgroup standard deviation when 1 < subgroup size <= 8.

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

SBAR

SBAR and POOLED are mutually exclusive subcommands that control the estimation of the within-subgroup standard deviation when subgroup size > 8.

SBAR specifies the average of the subgroup standard deviations to estimate standard deviation.

POOLED

SBAR and POOLED are mutually exclusive subcommands that control the estimation of the within-subgroup standard deviation when subgroup size > 8.

POOLED specifies the pooled standard deviation to estimate standard deviation. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups. Minitab displays an S chart in the output.

AMR

Specifies the average moving range to estimate the between subgroup standard deviation.

MMR

Specifies the median moving range to estimate the between subgroup standard deviation.

SRMSSD

Specifies the square root of half the mean of the squared successive differences to estimate the between subgroup standard deviation.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations.

RSPAN K

Specifies the length of the moving range. By default, a span of 2 is used because consecutive values have the greatest chance of being alike. The span must be < 100. RSPAN is ignored when the subgroup size > 1.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, SLIMITS 2 draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOLOWER K

Sets lower control limit bound for I chart.

LOUPPER K

Sets upper control limit bound for I chart.

RSLOWER K

Sets lower control limit bound for R or S chart.

RSUPPER K

Sets upper control limit bound for R or S chart.

VALOWER K
Sets lower control limit bound for MR chart.

VAUPPER K
Sets upper control limit bound for MR chart.

Tests

TEST K...K

TEST C
Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K
Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

Note RCHART, SCHART, and MRCHART only support tests 1 through 4.

DEFTEST applies to any chart that supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal l for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores standard deviations, one row for each historical stage.

BVARIATION C

Stores between standard deviations, one row for each historical stage.

WVARIATION C

Stores within standard deviation, one row for each historical stage.

TVARIATION C

Stores total standard deviation, one row for each historical stage.

PPOINTS C C C

Stores plotted points, one row for each subgroup.

CENLINE C C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C C C

Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

XBARCHART: Session command for creating an Xbar chart

XBARCHART [C...C *E*]
Creates an Xbar chart for subgroup means with subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The mean of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation is estimated by a pooled standard deviation by default. You can optionally enter an historical standard deviation with the SIGMA subcommand. Base your estimate on the average of the subgroup ranges with the RBAR subcommand. Or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

You can run the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

SRMSSD

Specifies the square root of MSSD to estimate σ.

BIASED

Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100. RSPAN is ignored when the subgroup size is less than 1.

When OMIT or ESTIMATE are used with RSPAN, any moving ranges which include omitted data are excluded from the calculations.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and

below the center line. SLIMITS 1 2 3 gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, SUBGROUP 6 specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, TEST 1 3 specifies perform tests 1 and 3. TEST 1:8 specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Note For Xbar charts and I charts, with extended control limits, only Test 1 is available. Test 1 fails when a point is outside the extended control limits.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use	1–6	3

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
	extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.		
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX C

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an

exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Extended Limits

EXTENDEDLIMITS

EXTENDEDLIMITS calculates control limits based on an extended limits method. If you do not specify a subcommand for the calculation method, then Minitab uses `BWITHIN`. You can use either `BWITHIN` or `MINMAX`, but not both.

BWITHIN [K1][K2]

Use the between/within method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (α-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the false alarm probability for the between-subgroup standard deviation (α-between). Specify a numeric value between 0 and 1. The default value is approximately `0.1336144`.

MINMAX [K1][K2]

Use the min/max method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (α-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the number of smallest and largest subgroup means to use to estimate the minimum and maximum process locations. Specify an integer from 1 through the number of subgroups. The default is `3`.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

RCHART: Session command for creating an R chart

RCHART [C...C *E*]
Creates an R chart for subgroup ranges, for subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The range of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation (sigma) can be estimated two ways. The default method for RCHART is to base the estimate on the average of the subgroup ranges with the RBAR subcommand. You can alternatively use the pooled standard deviation or enter an historical value with the SIGMA subcommand.

You can run the first four of the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

RSUB C...C

Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

SIGMA K

For variables control charts, the process is assumed to produce data from a stable population that often follows a normal distribution. The mean and standard deviation of a population distribution are called m and s, respectively. MU and SIGMA can be used to specify values for these parameters. The values you enter may be either actual known process parameters, or they may be estimates obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If m and s are not specified, they are estimated from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you

do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

Note RCHART supports only tests 1 through 4.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL
If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K
Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

SCHART: Session command for creating an S chart

SCHART [C...C *E*]

Creates a control chart for subgroup standard deviations, with subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line, and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The standard deviation of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation (sigma) can be estimated two ways. The default method uses the average of the subgroup standard deviations with the SBAR subcommand. You can alternatively base the estimate on the pooled standard deviation or enter an historical standard deviation with the SIGMA subcommand.

You can run the first four of the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

RSUB C...C

Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

SIGMA K

For variables control charts, the process is assumed to produce data from a stable population that often follows a normal distribution. The mean and standard deviation of a population distribution are called μ and σ, respectively. SIGMA can be used to specify values for these parameters. The values you enter may be either actual known process parameters, or they may be estimates obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ and σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

Note RCHART supports only tests 1 through 4.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C

Stores center line values, one row for each plotted point.

CONLIMITS C C

Stores control limit values, one row for each plotted point.

STAGES C

Stores stage indicators, one row for each plotted point.

SAMPSIZE C

Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C

Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

ZONE: Session command for creating a zone chart

ZONE [C...C *E*]

Creates a zone chart, with subgroup size = E.

ZONE can be used to plot a chart based on subgroup means or individual observations. The command is entered different ways, depending on the structure of your data. For more information, go to How to enter subgroup data on page 1246.

When the subgroup size is greater than one, the mean of all the observations in each subgroup is calculated. A point is plotted on the zone chart for each subgroup mean. Minitab generates a standardized zone chart when subgroup sizes are unequal. The process standard deviation (sigma) is estimated by a pooled standard deviation. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand. Base your estimate on the average of subgroup ranges with the RBAR subcommand. Or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

When the subgroup size is equal to one, a point is plotted on the zone chart for each individual observation. Sigma is estimated using a moving range of length two. You can optionally estimate sigma based on the median of the moving range with the MMR subcommand. Or specify a moving range of a different length with the RSPAN subcommand. Enter historical values of sigma with the SIGMA subcommand.

RSUB C...C

Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

SRMSSD

Specifies the square root of half the mean of the squared successive differences to estimate σ when subgroup size = 1.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating σ involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100. RSPAN is ignored when the subgroup size is less than 1.

Weights/Reset

RESET

Resets the cumulative score to zero following each out of control signal. When a process goes out of control, an attempt should be made to find and eliminate the cause of the problem. If the problem has been corrected, the cumulative score should be reset to zero.

WEIGHT K K K K

Use WEIGHT to specify the weights or scores assigned to points in each zone. For example, `WEIGHT 0 3 6 9` assigns a 0 to Zone 1, 3 to Zone 2, 6 to Zone 3, and 9 to Zone 4. The weight assigned to Zone 4 is also used as the critical value for determining when a process is out of control. In this case, the critical value would be 9. If you do not use WEIGHT, the default scores are 0, 2, 4, and 8.

Weights must be => 0, but do not need to be integers. When non-integer weights are used, ZONE rounds the scores displayed on the chart. Store the exact scores with SCORES.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX C

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K
Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Variables Charts for Individuals Data

IMRCHART: Session command for creating an I-MR chart

IMRCHART C...C

Draws an I chart (individual observations) and an MR chart (moving range) chart in the same graph.

Enter a column of individual observations. The chart for individual observations is in the upper half of the graph. The chart for moving ranges is in the lower half. Seeing both charts together allows you to examine both process level and process variation at the same time.

The estimate of sigma is based on MR / d2, the average of the moving range divided by an unbiasing constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. Use the RSPAN subcommand to change the length of the moving range. Alternatively, use the MMR subcommand to estimate sigma using the median of the moving range. Enter historical values of sigma with the SIGMA subcommand.

You can run all eight tests for special causes for the I portion of the chart. Only the first four tests are used with the MR portion. You can adjust the sensitivity of the tests with the DEFTEST subcommand.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100. RSPAN is ignored when the subgroup size is less than 1.

When OMIT or ESTIMATE are used with RSPAN, any moving ranges which include omitted data are excluded from the calculations.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOLOWER K

Sets lower control limit bound for I chart.

LOUPPER K

Sets upper control limit bound for I chart.

VALOWER K

Sets lower control limit bound for MR chart.

VAUPPER K

Sets upper control limit bound for MR chart.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Note For Xbar charts and I charts, with extended control limits, only Test 1 is available. Test 1 fails when a point is outside the extended control limits.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Extended Limits

EXTENDEDLIMITS

EXTENDEDLIMITS calculates control limits based on an extended limits method. If you do not specify a subcommand for the calculation method, then Minitab uses `BWITHIN`. You can use either `BWITHIN` or `MINMAX`, but not both.

BWITHIN [K1][K2]

Use the between/within method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the false alarm probability for the between-subgroup standard deviation (alpha-between). Specify a numeric value between 0 and 1. The default value is approximately `0.1336144`.

MINMAX [K1][K2]

Use the min/max method to calculate extended control limits.

- `K1` specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately `0.0026998`.
- `K2` specifies the number of smallest and largest subgroup means to use to estimate the minimum and maximum process locations. Specify an integer from 1 through the number of subgroups. The default is `3`.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C

Stores control limit values, one row for each plotted point.

STAGES C

Stores stage indicators, one row for each plotted point.

TRESULTS C...C

Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

ZMRCHART: Session command for creating a Z-MR chart

ZMRCHART C...C *E*

Generates control charts for standardized individual observations (Z) and moving ranges (MR) on measurement data from a short run process, with part indicators = E.

ZMRCHART displays a chart for standardized individual observations (Z) above a chart for moving ranges (MR).

To create a Z-MR chart, type ZMRCHART followed by a data column and a column that contains the part/product name or number for each measurement. The columns must have the same number of rows. Thus, each row contains a measurement, and a part or product name. The part/product name column can contain text or numeric data. The part/product data defines the groupings for estimating process parameters. In addition, each time Minitab encounters a change in the part/product name column a new run is defined. You can use the SET on page 81 command to enter patterned data for the part/product number. For more information, go to Entering patterned data for the SET session command on page 1226.

ZMRCHART standardizes the measurement data by subtracting the mean and dividing the result by the standard deviation. Standardizing allows data collected from different runs to be evaluated by interpreting a single control chart. You can estimate the means and standard deviations from the data, or supply historical values using the CENTER and SPREAD subcommands.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

CENTER C...C

CENTER K...K

Use CENTER to enter historical values to standardize the data. The data is standardized by subtracting the mean—to center the data—and dividing by the standard deviation. When you use CENTER, Minitab uses the data given with the column or constant argument to standardize the data rather than estimates obtained from the measurement data. You must enter a value for each observation in the data column. Thus, the CENTER column must be the same length as the data column.

You can use CENTER to enter a historical mean for each part/product. Using historical means allows you to compare your process with past performance. When you use known means to center the data, the chart reflects whether the process is functioning as it has in the past with respect to location. That is, the chart shows whether each part/product mean is the same as previously established. When the process is functioning the same in respect to location, the values are distributed (equally) around the center line.

You can also use nominal specifications—target values—for each part/product to center the data. Using nominal specifications to center the data allows you to compare your process to desired performance. The nominal specifications are established target values for the dimension of interest for each part/product. When you use nominal specifications to center the data, the chart reflects whether the process is producing parts/products which are on target, or whether the process is biased.

SPREAD C...C

SPREAD K...K

Use SPREAD to use historical values to estimate standard deviation.

Estimate

SEPARATE

SEPARATE estimates standard deviation for each run independently.

POOLED

POOLED combines the data from runs of the same part together, then estimates standard deviation for that part. This is repeated to estimate standard deviation for all parts.

EQUAL

EQUAL pools all the data—across runs and parts—to obtain a common estimate of standard deviation.

PROP

PROP takes the natural log of the data, then pools the transformed data--across all runs and all parts--to obtain a common estimate of standard deviation for the transformed data. The natural log transformation stabilizes the variation in cases where variation increases as the size of the measurement increases.

AMR

Use to base the standard deviation estimate on the average of the moving range divided by an unbiasing constant.

MMR

Use to base the standard deviation estimate on the median of the moving range.

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. K must be < 100.

Split

SPLIT [K]

Split series by runs (when K is not specified) or by subset (when K is specified).

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

VALOWER K

Sets lower control limit bound for MR chart.

VAUPPER K

Sets upper control limit bound for MR chart.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

ZMRCHART supports all 8 tests.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K
Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

Storage

LOCATION C
Stores means.

VARIATION C
Stores standard deviations.

PPOINTS C C
Stores plotted points.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

ICHART: Session command for creating an I chart

ICHART C...C
Draws a control chart for individual observations.

Enter a column of individual observations. ICHART plots the individual observations on a chart.

The estimate of sigma is based on MR / d2, the average of the moving range divided by an unbiasing constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN to change the length of the moving range. Alternatively, use the MMR subcommand or the SRMSSD subcommand to estimate the standard deviation or enter historical standard deviations with the SIGMA subcommand.

You can run the eight common tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the tests, use the DEFTEST subcommand.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

AMR

Specifies the average moving range to estimate σ.

MMR

Specifies the median moving range to estimate σ.

SRMSSD

Specifies the square root of MSSD to estimate σ.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (SRMSSD).

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. K must be < 100.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bound. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bound. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Note For Xbar charts and I charts, with extended control limits, only Test 1 is available. Test 1 fails when a point is outside the extended control limits.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Extended Limits

EXTENDEDLIMITS

EXTENDEDLIMITS calculates control limits based on an extended limits method. If you do not specify a subcommand for the calculation method, then Minitab uses BWITHIN. You can use either BWITHIN or MINMAX, but not both.

BWITHIN [K1][K2]

Use the between/within method to calculate extended control limits.

- K1 specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately 0.0026998.
- K2 specifies the false alarm probability for the between-subgroup standard deviation (alpha-between). Specify a numeric value between 0 and 1. The default value is approximately 0.1336144.

MINMAX [K1][K2]

Use the min/max method to calculate extended control limits.

- K1 specifies the false alarm probability for the within-subgroup standard deviation (alpha-within). Specify a numeric value between 0 and 1. The default value gives the same false alarm rate as a traditional 3-sigma Shewhart chart for an in-control process, approximately 0.0026998.
- K2 specifies the number of smallest and largest subgroup means to use to estimate the minimum and maximum process locations. Specify an integer from 1 through the number of subgroups. The default is 3.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

MRCHART: Session command for creating an MR chart

MRCHART C...C
Draws an MR (moving range) control chart.

Enter a column of individual observations. The moving ranges of length K are calculated for all nonmissing data and plotted on the chart.

The estimate of standard deviation is based on MR / d2, the average of the moving range divided by an unbiasing constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. Use the RSPAN subcommand to change the length of the moving range. Alternatively, use the MMR subcommand to estimate the standard deviation using the median of the moving range. Enter historical standard deviations with the SIGMA subcommand.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

AMR

Specifies the average moving range to estimate σ.

MMR

Specifies the median moving range to estimate σ.

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. K must be < 100.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Acceptance Control Charts

AXRCHART: Session command for creating an Acceptance Xbar-R chart

AXRCHART [C...C *E*]
Creates an Acceptance Xbar-R chart for subgroup means and ranges with subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The mean of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation is estimated by the average of subgroup ranges. You can optionally enter an historical standard deviation with the SIGMA subcommand. Base your estimate on the pooled standard deviation with the POOLED subcommand.

You can run tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the R chart tests, use the DEFTEST subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Specifications and acceptance criteria

LSPEC K...K

LSPEC C
Specifies the lower specification limit. Specify one constant to use the same lower specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of `LSPEC` or `USPEC` is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

USPEC K...K

USPEC C

Specifies the upper specification limit. Specify one constant to use the same upper specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

POSPEC K

Specifies the accepted percentage outside of each specification limit. K must be greater than 0 and less than 100. This subcommand is required.

PAINTERVENTION K

Specifies the probability of intervention when the percentage outside a specification limit exceeds the accepted percentage. K must be greater than 0 and less than 1. The default value is 0.99.

Note When you use a Box-Cox transformation, Minitab transforms the values that you enter for LSPEC and USPEC for the calculations.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

BIASED

Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

VALOWER K

Sets lower control limit bound for R chart.

VAUPPER K

Sets upper control limit bound for R chart.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the R chart tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. You can specify tests 1 through 4. For example, `TEST 1 3` performs tests 1 and 3. `TEST 1:4` performs all four tests. TEST with a column is useful when writing macros.

For the acceptance chart, Test 1 identifies points outside the acceptance chart control limits. For the R chart, tests 1 through 4 apply.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the R chart tests for special causes.

The test definitions only apply to this chart.

DEFTEST applies only to the R chart. The first K in each pair corresponds to the number of the test; the second K defines the test. The first K can be 1 through 4. If you specify `DEFTEST 1 K`, the definition applies only to Test 1 for the R chart.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (R chart)	1-6	3
2	K points in a row on same side of center line (R chart)	7-11	9
3	K points in a row, all increasing or all decreasing (R chart)	5-8	6
4	K points in a row, alternating up and down (R chart)	12-14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an

exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

LSL C...C
Stores lower specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify a lower specification limit. When you use a Box-Cox transformation, Minitab stores the lower specification limits in transformed units.

USL C...C
Stores upper specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify an upper specification limit. When you use a Box-Cox transformation, Minitab stores the upper specification limits in transformed units.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

AXSCHART: Session command for creating an Acceptance Xbar-S chart

AXSCHART [C...C *E*]
Creates an Acceptance Xbar-S chart for subgroup means and subgroup standard deviations with subgroup size = E.

If the subgroups are in columns, enter data with the CSUB subcommand. If the subgroups are in rows, enter data with the RSUB subcommand. For details, go to How to enter subgroup data on page 1246. The X chart is drawn in the upper half of the screen; the S chart in the lower half. Seeing both charts together allows you to examine both process level and process variation at the same time.

The process standard deviation is estimated the average of the subgroup standard deviations. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand, or base your estimate on the pooled standard deviations with the POOLED subcommand.

You can run tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the S chart tests, use the DEFTEST subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Specifications and acceptance criteria

LSPEC K...K

LSPEC C

Specifies the lower specification limit. Specify one constant to use the same lower specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

USPEC K...K

USPEC C

Specifies the upper specification limit. Specify one constant to use the same upper specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

POSPEC K

Specifies the accepted percentage outside of each specification limit. K must be greater than 0 and less than 100. This subcommand is required.

PAINTERVENTION K

Specifies the probability of intervention when the percentage outside a specification limit exceeds the accepted percentage. K must be greater than 0 and less than 1. The default value is 0.99.

Note When you use a Box-Cox transformation, Minitab transforms the values that you enter for LSPEC and USPEC for the calculations.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, the column can contain different values for each historical group. If you enter one constant, all historical stages use that value. Otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, the column can contain different values for each historical group. If you enter one constant, all historical stages use that value. Otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED
Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

SBAR
Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

BIASED
Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED
Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

S limits

SLIMITS K...K

SLIMITS C
Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

VALOWER K
Sets lower control limit bound for S chart.

VAUPPER K
Sets upper control limit bound for S chart.

SUBGROUP K
If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C
Performs one or more tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the S chart tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. You can specify tests 1 through 4. For example, `TEST 1 3` performs tests 1 and 3. `TEST 1:4` performs all four tests. TEST with a column is useful when writing macros.

For the acceptance chart, Test 1 identifies points outside the acceptance chart control limits. For the S chart, tests 1 through 4 apply.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the S chart tests for special causes.

The test definitions only apply to this chart.

DEFTEST applies only to the S chart. The first K in each pair corresponds to the number of the test; the second K defines the test. The first K can be 1 through 4. If you specify `DEFTEST 1 K`, the definition applies only to Test 1 for the S chart.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (S chart)	1-6	3
2	K points in a row on same side of center line (S chart)	7-11	9
3	K points in a row, all increasing or all decreasing (S chart)	5-8	6
4	K points in a row, alternating up and down (S chart)	12-14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal l for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C C
Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

LSL C...C
Stores lower specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify a lower specification limit. When you use a Box-Cox transformation, Minitab stores the lower specification limits in transformed units.

USL C...C
Stores upper specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify an upper specification limit. When you use a Box-Cox transformation, Minitab stores the upper specification limits in transformed units.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

AXBARCHART: Session command for creating an Acceptance Xbar chart

AXBARCHART [C...C *E*]
Creates an Acceptance Xbar chart for subgroup means with subgroup size = E.

If the subgroups are in columns, enter a data column followed by a subgroup indicator. If the subgroups are in rows, omit the arguments on the main command line and use the RSUB subcommand to enter the data. For details, go to How to enter subgroup data on page 1246. The mean of the data in each subgroup is calculated and plotted on the chart.

The process standard deviation is estimated by a pooled standard deviation by default. You can optionally enter an historical standard deviation with the SIGMA subcommand. Base your estimate on the average of the subgroup ranges with the RBAR subcommand. Or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

You can run Test 1 for special causes. Use the TEST subcommand to specify whether to perform the test.

RSUB C...C
Specifies that data are in subgroups across rows.

Specifications and acceptance criteria

LSPEC K...K

LSPEC C
Specifies the lower specification limit. Specify one constant to use the same lower specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

USPEC K...K

USPEC C
Specifies the upper specification limit. Specify one constant to use the same upper specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

POSPEC K
Specifies the accepted percentage outside of each specification limit. K must be greater than 0 and less than 100. This subcommand is required.

PAINTERVENTION K
Specifies the probability of intervention when the percentage outside a specification limit exceeds the accepted percentage. K must be greater than 0 and less than 1. The default value is 0.99.

Note When you use a Box-Cox transformation, Minitab transforms the values that you enter for LSPEC and USPEC for the calculations.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

SRMSSD

Specifies the square root of MSSD to estimate σ.

BIASED

Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100. RSPAN is ignored when the subgroup size is less than 1.

When OMIT or ESTIMATE are used with RSPAN, any moving ranges which include omitted data are excluded from the calculations.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and

below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K

TEST C

Specifies whether to perform Test 1 for special causes. Enter 0 or 1, or a column that contains 0 or 1. Enter 1 to identify points outside the acceptance chart control limits. Enter 0 to perform no tests and override any set control chart options.

When a point fails Test 1, it is marked with the test number on the chart.

Minitab will perform the test when subgroup sizes are unequal.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX C

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

LSL C...C
Stores lower specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify a lower specification limit. When you use a Box-Cox transformation, Minitab stores the lower specification limits in transformed units.

USL C...C
Stores upper specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify an upper specification limit. When you use a Box-Cox transformation, Minitab stores the upper specification limits in transformed units.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

AIMRCHART: Session command for creating an Acceptance I-MR chart

AIMRCHART C...C
Draws an Acceptance I chart (individual observations) and an MR chart (moving range) in the same graph.

Enter a column of individual observations. The chart for individual observations is in the upper half of the graph. The chart for moving ranges is in the lower half. Seeing both charts together allows you to examine both process level and process variation at the same time.

The estimate of sigma is based on MR / d2, the average of the moving range divided by an unbiasing constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. Use the RSPAN subcommand to change the length of the moving range. Alternatively, use the MMR subcommand to estimate sigma using the median of the moving range. Enter historical values of sigma with the SIGMA subcommand.

You can run tests for special causes. Use the TEST subcommand to specify which tests to use. To adjust the sensitivity of the MR chart tests, use the DEFTEST subcommand.

Specifications and acceptance criteria

LSPEC K...K

LSPEC C

Specifies the lower specification limit. Specify one constant to use the same lower specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

USPEC K...K

USPEC C

Specifies the upper specification limit. Specify one constant to use the same upper specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

POSPEC K

Specifies the accepted percentage outside of each specification limit. K must be greater than 0 and less than 100. This subcommand is required.

PAINTERVENTION K

Specifies the probability of intervention when the percentage outside a specification limit exceeds the accepted percentage. K must be greater than 0 and less than 1. The default value is 0.99.

Note When you use a Box-Cox transformation, Minitab transforms the values that you enter for LSPEC and USPEC for the calculations.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR
Specifies the median moving range to estimate σ when subgroup size = 1.

BIASED
Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED
Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K
RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100. RSPAN is ignored when the subgroup size is less than 1.

When OMIT or ESTIMATE are used with RSPAN, any moving ranges which include omitted data are excluded from the calculations.

S limits

SLIMITS K...K

SLIMITS C
Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

VALOWER K
Sets lower control limit bound for MR chart.

VAUPPER K
Sets upper control limit bound for MR chart.

Tests

TEST K...K

TEST C
Performs one or more tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the MR chart tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. You can specify tests 1 through 4. For example, `TEST 1 3` performs tests 1 and 3. `TEST 1:4` performs all four tests. TEST with a column is useful when writing macros.

For the acceptance chart, Test 1 identifies points outside the acceptance chart control limits. For the MR chart, tests 1 through 4 apply.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the MR chart tests for special causes.

The test definitions only apply to this chart.

DEFTEST applies only to the MR chart. The first K in each pair corresponds to the number of the test; the second K defines the test. The first K can be 1 through 4. If you specify `DEFTEST 1 K`, the definition applies only to Test 1 for the MR chart.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (MR chart)	1-6	3
2	K points in a row on same side of center line (MR chart)	7-11	9
3	K points in a row, all increasing or all decreasing (MR chart)	5-8	6
4	K points in a row, alternating up and down (MR chart)	12-14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C C
Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

LSL C...C
Stores lower specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify a lower specification limit. When you use a Box-Cox transformation, Minitab stores the lower specification limits in transformed units.

USL C...C
Stores upper specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify an upper specification limit. When you use a Box-Cox transformation, Minitab stores the upper specification limits in transformed units.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

AICHART: Session command for creating an Acceptance I chart

AICHART C...C
Draws an Acceptance I chart for individual observations.

Enter a column of individual observations. AICHART plots the individual observations on a chart.

The estimate of sigma is based on MR / d2, the average of the moving range divided by an unbiasing constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN to change the length of the moving range. Alternatively, use the MMR subcommand or the SRMSSD subcommand to estimate the standard deviation or enter historical standard deviations with the SIGMA subcommand.

You can run Test 1 for special causes. Use the TEST subcommand to specify whether to perform the test.

Specifications and acceptance criteria

LSPEC K...K

LSPEC C

Specifies the lower specification limit. Specify one constant to use the same lower specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

USPEC K...K

USPEC C

Specifies the upper specification limit. Specify one constant to use the same upper specification limit for all variables, or specify one constant for each variable. You can also specify a column that contains one specification limit for each variable.

At least one of LSPEC or USPEC is required. If you specify both, each lower specification limit must be less than the corresponding upper specification limit.

POSPEC K

Specifies the accepted percentage outside of each specification limit. K must be greater than 0 and less than 100. This subcommand is required.

PAINTERVENTION K

Specifies the probability of intervention when the percentage outside a specification limit exceeds the accepted percentage. K must be greater than 0 and less than 1. The default value is 0.99.

Note When you use a Box-Cox transformation, Minitab transforms the values that you enter for LSPEC and USPEC for the calculations.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

AMR

Specifies the average moving range to estimate σ.

MMR
Specifies the median moving range to estimate σ.

SRMSSD
Specifies the square root of MSSD to estimate σ.

BIASED
Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED
Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (SRMSSD).

RSPAN K
RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. K must be < 100.

S limits

SLIMITS K...K

SLIMITS C
Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

Tests

TEST K

TEST C
Specifies whether to perform Test 1 for special causes. Enter 0 or 1, or a column that contains 0 or 1. Enter 1 to identify points outside the acceptance chart control limits. Enter 0 to perform no tests and override any set control chart options.

When a point fails Test 1, it is marked with the test number on the chart.

Stages

HPROCESS C
HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a

control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores standard deviations, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C

Stores center line values, one row for each plotted point.

CONLIMITS C C

Stores control limit values, one row for each plotted point.

STAGES C

Stores stage indicators, one row for each plotted point.

LSL C...C

Stores lower specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify a lower specification limit. When you use a Box-Cox transformation, Minitab stores the lower specification limits in transformed units.

USL C...C

Stores upper specification limits. The number of columns equals the number of models. Missing values are included in the column if you do not specify an upper specification limit. When you use a Box-Cox transformation, Minitab stores the upper specification limits in transformed units.

TRESULTS C...C

Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Attributes Control Charts

PDIAGNOSTIC: Session command for determining whether to use a P chart or a Laney P' chart

PDIAGNOSTIC C *E*

Use this test to determine whether a traditional P chart (PCHART on page 544) or a Laney P' chart (PPRIME on page 549) is more appropriate to chart the proportion of defectives in your sample. If the test determines that your data exhibit overdispersion or underdispersion a Laney P' chart may more accurately distinguish between common cause variation and special cause variation.

For more information and an example of data entry, go to How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART on page 1244.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

PCHART: Session command for creating a P chart

PCHART C *E*

Draws a chart for proportion of defectives.

Each entry in the column is the number of defectives for one sample. Minitab assumes the sample came from a binomial distribution with parameters n and p. E is the subgroup size for the sample, which can be a column (if the subgroup size varies) or a constant (if all samples are the same size). Minitab plots the proportion of defectives, X / n, for each sample. For more information and an example of data entry, go to How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART on page 1244.

PCHART estimates the process proportion defective from the data by default. The process proportion is the center line on the chart and is also used to calculated the control limits. If you have a historical value for the process proportion defective, use the P subcommand to enter it.

The center line and control limits vary with the subgroup size in PCHART. Use the SUBGROUP subcommand to make them constants.

You can run four common tests for special causes for PCHART. Use the TEST subcommand to specify which tests to use.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

P K...K

P C

For PCHART and NPCHART, the process is assumed to produce data from a population that follows a binomial distribution. The parameters of the binomial distribution are the sample size, n, and the proportion of defectives, p. P can be used to specify a value for the process proportion defective. The value you use may be either an actual known process proportion, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value; otherwise, enter a value for each historical group.

If P is not specified, it is estimated by the overall proportion of defectives in the sample.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores proportions, one row for each historical group.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C

Stores center line values, one row for each subgroup.

CONLIMITS C C
Stores control limit values, one row for each subgroup.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

PPRIMECHART: Session command for creating a Laney P' chart

PPRIMECHART C *E*
Draws a Laney P' chart for proportion of defectives. The Laney P' chart is similar to a traditional P chart (PCHART on page 544). Both charts help you to monitor the proportion of defective items that are produced by your process. The Laney P' chart can be useful in the following situations:

- You have large subgroups and your data exhibit overdispersion.

  Overdispersion can cause the points on a traditional P chart to appear to be out of control when they are not. For the Laney P' chart, the definition of common cause variation includes not only the within-subgroup variation, but also the average variation between consecutive subgroups. If there is overdispersion, the control limits on a Laney P' chart are wider than those of a traditional P chart. The wider control limits mean that only important deviations in your process are identified as out of control.

- Your data exhibit underdispersion.

  Underdispersion, which can occur with subgroups of any size, is often caused by a lack of randomness. Underdispersion can result in control limits that are too wide for the data. The Laney P' chart corrects for underdispersion by calculating narrower control limits.

You can use P chart diagnostic (PDIAGNOSTIC on page 543) to test for overdispersion and underdispersion.

The calculations for the Laney P' chart include Sigma Z, which is an adjustment for overdispersion or underdispersion. A Sigma Z value of 1 indicates that no adjustment is necessary and that the Laney P' chart is exactly the same as a traditional P chart.

For more information and an example of data entry, go to How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART on page 1244.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

P K...K

P C

The process is assumed to produce data from a population that follows a binomial distribution. The parameters of the binomial distribution are the sample size, n, and the proportion of defectives, p. P can be used to specify a value for the process proportion defective. The value you use may be either an actual known process proportion, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value; otherwise, enter a value for each historical group.

If P is not specified, it is estimated by the overall proportion of defectives in the sample.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bound. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bound. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores proportions, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C

Stores center line values, one row for each subgroup.

CONLIMITS C C

Stores control limit values, one row for each subgroup.

STAGES C

Stores stage indicators, one row for each plotted point.

SAMPSIZE C

Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C

Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

NPCHART: Session command for creating a chart for the number of defectives

NPCHART C *E*

Draws a chart for the number of defectives.

Each entry in the column is the number of defectives for one sample. The sample is assumed to have come from a binomial distribution with parameters n and p. E is the subgroup size for the sample, which can be a column (if the subgroup size varies) or a constant (if all samples are the same size). Minitab plots the number of defectives for each sample. For more information and an example of data entry, go to How to enter data for NPCHART on page 1243.

NPCHART estimates the process proportion defective from the data by default. The process proportion is the center line on the NPCHART and is also used to the control limits. If you have a historical value for the process proportion defective, use the P subcommand to enter it.

The center line and control limits vary with the subgroup size in NPCHART. Use the SUBGROUP subcommand to make them constants.

You can run four common tests for special causes for NPCHART. Use the TEST subcommand to specify which tests to use.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

P K...K

P C

For NPCHART, the process is assumed to produce data from a population that follows a binomial distribution. The parameters of the binomial distribution are the sample size, n, and the proportion of defectives, p. P can be used to specify a value for the process proportion defective. The value you use may be either an actual known process proportion, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value; otherwise, enter a value for each historical group.

If P is not specified, it is estimated by the overall proportion of defectives in the sample.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K
Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores proportions, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

UPRIMECHART: Session command for creating a Laney U' chart

UPRIMECHART C *E*
Draws a Laney U' chart for proportion of defectives. The Laney U' chart is similar to a traditional U chart (UCHART on page 565). Both charts help you to monitor the number of defects per unit that are produced by your process. The Laney U' chart can be useful in the following situations:

- You have large subgroups and your data exhibit overdispersion.

  Overdispersion can cause the points on a traditional U chart to appear to be out of control when they are not. For the Laney U' chart, the definition of common cause variation includes not only the within-subgroup variation, but also the average variation between consecutive subgroups. If there is overdispersion, the control limits on a Laney U' chart are wider than those of a traditional U chart. The wider control limits mean that only important deviations in your process are identified as out of control.

- Your data exhibit underdispersion.

  Underdispersion, which can occur with subgroups of any size, is often caused by a lack of randomness. Underdispersion can result in control limits that are too wide for the data. The Laney U' chart corrects for underdispersion by calculating narrower control limits.

You can use U chart diagnostic (UDIAGNOSTIC on page 570) to test for overdispersion and underdispersion.

The calculations for the Laney U' chart include Sigma Z, which is an adjustment for overdispersion or underdispersion. A Sigma Z value of 1 indicates that no adjustment is necessary and that the Laney U' chart is exactly the same as a traditional U chart.

For more information and an example of data entry, go to How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART on page 1246.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

The process is assumed to produce data from a population that follows a Poisson distribution. The mean and variance of this distribution are both denoted by μ. MU can be used to specify a value for this parameter. The value you enter may be either an actual known process mean, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value. Otherwise, enter a value for each historical group.

If μ is not specified, it is estimated by the overall sample mean.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bound. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bound. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores estimates of the means, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CENLINE C

Stores center line values, one row for each subgroup.

CONLIMITS C C
Stores control limit values, one row for each subgroup.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

UCHART: Session command for creating a U chart

UCHART C *E*
Draws a chart of the number of defects per unit sampled.

Each entry in the specified column is the number of defects for one sample. Minitab assumes the sample comes from a Poisson distribution with the parameter m. E is the sample size for each sample in the column, which can be a column (if the subgroup size varies) or a constant (if all samples are the same size). Minitab plots the number of defects per unit, or X / n, for each sample.

UCHART estimates the process average number of defects from the data by default. The process average is the center line on a UCHART and is also used to calculated the control limits. If you have an historical value for the process average number of defects, use the MU subcommand to enter it.

The control limits vary with the subgroup size in UCHART. Use the SUBGROUP command to make them constants. For examples, go to How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART on page 1246.

You can run four common tests for special causes for UCHART. Use the TEST subcommand to specify which tests to use.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

P C

For CCHART and UCHART, the process is assumed to produce data from a population that follows a Poisson distribution. The mean and variance of this distribution are both denoted by μ. MU can be used to specify a value for this parameter. The value you enter may be either an actual known process mean, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value. Otherwise, enter a value for each historical group.

If μ is not specified, it is estimated by the overall sample mean.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores estimates of the means, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each subgroup.

CONLIMITS C C
Stores control limit values, one row for each subgroup.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

UDIAGNOSTIC: Session command for determining whether to use a U chart or a Laney U' chart

UDIAGNOSTIC C *E*
Use this test to determine whether a traditional U chart (UCHART on page 565) or a Laney U' chart (UPRIMECHART on page 559) is more appropriate to chart the number of defects per unit in your sample. If the test determines that your data exhibit overdispersion or underdispersion a Laney U' chart may more accurately distinguish between common cause variation and special cause variation.

For more information and an example of data entry, go to How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART on page 1246.

Estimate

OMIT K...K

OMIT C
Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup

32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

CCHART: Session command for creating a C chart

CCHART C

Draws a chart of the number of defects.

Each entry in the specified column contains the number of defects for one sample. The sample is assumed to have come from a Poisson distribution with parameter m. Minitab plots the number of defects, or X, for each sample. For more information and an example of data entry, go to How to enter data for CCHART on page 1242.

CCHART estimates the process average number of defects from the data by default. The process average is the center line on the CCHART and is also used to calculated the control limits. Use the MU subcommand to enter a historical value for the process average number of defects.

You can run four common tests for special causes for CCHART. Use the TEST subcommand to specify which tests to use.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C

For CCHART, the process is assumed to produce data from a population that follows a Poisson distribution. The mean and variance of this distribution are both denoted by μ. You can use MU to specify a value for this parameter. You can specify either an actual known process mean, or an estimate obtained from past data.

If you use a column, it can contain different values for each stage. If you enter one constant, all historical groups use that value. Otherwise, enter a value for each historical group.

If you don't specify μ, it is estimated by the overall sample mean.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart that supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores proportions, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Time-Weighted Control Charts

MACHART: Session command for creating a moving average chart

MACHART [C...C *E*]
Draws an MA (moving average) control chart.

MACHART can plot a chart based on either subgroup means or individual observations with subgroup size = E. When plotting a chart based on subgroup means, the command is entered two ways, depending on the structure of your data. For details, go to How to enter subgroup data on page 1246.

To plot a chart of individual observations, enter the column followed by a subgroup indicator of 1. For example, `MACHART C1 1` plots a chart of the individual observations contained in C1. To enter data for individual observations, you must have all the data from a process in a single column. Since there are no subgroups, you do not need a subgroup indicator column or a subgroup size.

If your data are in subgroups, then the mean of all the observations in each subgroup is calculated. Moving averages are then formed from these means. The process standard deviation (sigma) is estimated by a pooled standard deviation. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand, base your estimate on the average of subgroup ranges with the RBAR subcommand, or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

If your data are individual observations, then moving averages are formed from the individual observations. In this case, sigma is estimated using a moving range of length two. You can optionally estimate σ based on the median of the moving range with the MMR subcommand. Or specify a moving range of a different length with the RSPAN subcommand. Enter an historical value for the process standard deviation with the SIGMA subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you

do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

SRMSSD

Specifies the square root of half the mean of the squared successive differences to estimate σ when subgroup size = 1.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K

Specifies the length of the moving range. By default, a span of two is used because consecutive values have the greatest chance of being alike. The span must be < 100.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C
Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

EWMACHART: Session command for creating an EWMA chart

EWMACHART [C...C *E*]
Draws an EWMA (exponentially weighted moving average) control chart.

EWMACHART can be used to plot a chart based on either subgroup means or individual observations. When plotting a chart based on subgroup means, the command is entered two ways, depending on the structure of your data. For details, go to How to enter subgroup data on page 1246.

To plot a chart of individual observations, enter the column followed by a subgroup indicator of 1. For example, `EWMACHART C1 1` plots a chart of the individual observations contained in C1.

If your data are in subgroups, then the mean of all the observations in each subgroup is calculated. Exponentially weighted moving averages are then formed from these means. The process standard deviation (sigma) is estimated by a pooled standard deviation. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand, base your estimate on the average of subgroup ranges with the RBAR subcommand, or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

If your data are individual observations, then exponentially weighted moving averages are formed from the individual observations. In this case, sigma is estimated using a moving range of length two. You can optionally estimate σ based on the median of the moving range with the MMR subcommand. Or specify a moving range of a different length with the RSPAN subcommand. Enter an historical value for the process standard deviation with the SIGMA subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Weight

WEIGHT K

Specifies the weight used in the exponentially weighted moving average. The weight must be between 0 and 1. The default is 0.2. By changing the weight used and the number of sigmas for the control limits, you can construct a chart with very specific properties. Combinations of these two parameters are often chosen by using an ARL (average run length) table.

Parameters

MU K...K

MU C

Specify a value for mu (denoted as μ), the mean of the population. The value that you enter can be either an actual known process parameter or an estimate that was obtained from past data.

If you use a column, it can contain different values for each historical stage. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical stage.

If you do not specify a value, Minitab estimates it from the data.

SIGMA K...K

SIGMA C

Specify a value for sigma (denoted as σ), the standard deviation of the population. The value that you enter can be either an actual known process parameter or an estimate that was obtained from past data.

If you use a column, it can contain different values for each historical stage. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical stage.

If you do not specify a value, Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using EXCLUDE 1), then decide to omit observation 2 (using OMIT 2). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size > 1.

SRMSSD

Specifies the square root of half the mean of the squared successive differences to estimate σ when subgroup size > 1.

BIASED

Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with the pooled standard deviation method of estimating σ.

RSPAN K

Specifies the length of the moving range. By default, a span of two is used because consecutive values have the greatest chance of being alike. The span must be < 100.

S limits

SLIMITS K...K

SLIMITS C

Draws additional sigma limit lines above and below the mean at the standard deviation values that you specify. The arguments for SLIMITS are positive numbers. Each argument you give draws two horizontal lines, one above and one below the mean. For example, `SLIMITS 2` draws control limits at two standard deviations above and below the center line. `SLIMITS 1 2 3` gives three lines above and three lines below the center line at the first, second, and third standard deviations.

LOWER K

Sets lower control limit bounds. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bounds. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST K K...K K

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C
Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

CUSUM: Session command for creating a CUSUM chart

CUSUM [C...C *E*]
Draws a CUSUM (cumulative sum) control chart with subgroup size = E.

Use CUSUM to plot a chart based on subgroup means or individual observations. You can enter the command in different ways, depending on the structure of your data. For details, go to How to enter subgroup data on page 1246. When data are arranged across rows, you can optionally specify the number of models you would like. For example, if you have data across rows c1 - c6 and you have two models (C1–C3 and C4–C6), you can enter `RSUB C1–C6 2` to produce two control charts. The number of models you specify must divide evenly into the number of columns on the RSUB subcommand.

CUSUM plots two 1-sided CUSUMs. The upper CUSUM is for detecting upward shifts in the level of the process; the lower CUSUM for detecting downward shifts. This chart uses an UCL and an LCL to determine when an out-of-control situation has occurred.

Use VMASK on page 596 to plot a single 2-sided CUSUM. This chart uses a V-mask to determine when an out-of-control situation has occurred.

The CUSUM chart is defined by two parameters, h and k. These parameters are often referred to as a "CUSUM plan" and are specified on the PLAN subcommand. In the default chart, h is the number of standard deviations between the center line and the control limits, and k is the allowable "slack" in the process.

If your data are in subgroups, then Minitab calculates the mean of all the observations in each subgroup, then calculates the CUSUM statistics using the subgroup means. All subgroups must be the same size. Minitab estimates the process standard deviation by a pooled standard deviation. You can optionally enter an historical value for the process standard deviation with the SIGMA subcommand, base your estimate on the average of subgroup ranges with the RBAR subcommand, or base your estimate on the average of the subgroup standard deviations with the SBAR subcommand.

If your data are individual observations, then Minitab calculates the CUSUM statistics using the individual observations. Minitab estimates sigma using a moving range length of 2. You can optionally estimate sigma based on the median of the moving range with the MMR subcommand. You can also specify a moving range of a different length with the RSPAN subcommand. Enter historical standard deviation values with the SIGMA subcommand.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Target

TARGET K

Specifies a target. CUSUM statistics are cumulative deviations from this target, or nominal, specification.

Parameters

SIGMA K...K

SIGMA C

For variables control charts, the process is assumed to produce data from a stable population that often follows a normal distribution. You can use SIGMA to specify values for the standard deviation. You can enter either actual known process parameters, or estimates obtained from past data.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup

32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Use to base your estimate on the average of the subgroup standard deviations.

AMR

When plotting control charts for individual observations, the estimate of standard deviation is based on the average of the moving range divided by an unbiasing constant.

MMR

Use to base the estimate on the median of the moving range.

SRMSSD

The SRMSSD subcommand estimates the standard deviation using the square root of half the mean of the squared successive differences. Use this subcommand for charts which allow individual points where the standard deviation does not have to be estimated using the moving range.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating s involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K

RSPAN specifies the length of the moving range. By default, Minitab uses a moving range of length 2, since consecutive values have the greatest chance of being alike. Use RSPAN only when you want to change the length of the moving range from the default of 2. K must be < 100.

Plan/Type

PLAN K K

Specifies a CUSUM plan, which is defined by the parameters h and k. The first value entered is h, the second is k. Both h and k must be greater than 0. The default values are 4 and 0.5.

FIR K

Uses the FIR (fast initial response) method to initialize the one-sided CUSUMs. Normally, they are initialized at 0. If the process is out of control at startup, the CUSUMs will not detect the situation for several subgroups. The FIR method initializes the CUSUMs at a specified number of sigma's above and below 0. This method has been shown to reduce the number of subgroups needed to detect problems at startup.

RESET

Resets the CUSUMs to their initial values whenever an out-of-control signal is generated. When a process goes out of control, an attempt should be made to find and eliminate the cause of the problem. If the problem has been corrected, the CUSUMs should be reset to their initial values.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, `SUBGROUP 6` specifies to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, then Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C C
Stores plotted points, one row for each subgroup.

CENLINE C C
Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

VMASK: Session command for creating a two-sided CUSUM chart

VMASK [C...C *E*]
VMASK displays a single 2-sided CUSUM, with subgroup size = E. This chart uses a V-mask to determine when an out-of-control situation has occurred. Use CUSUM on page 590 to display a 1-sided CUSUM.

CUSUM charts are defined by two parameters, h and k. These parameters are often referred to as a "CUSUM plan" and are specified on the PLAN subcommand. In the VMASK chart, Minitab calculates the half-width of the V-mask (H) at the point of origination by H = h * s. The slope of the V-mask arms is defined by k. Values of h and k are often selected from ARL (average run length) tables.

RSUB C...C
Specifies that data are in subgroups across rows.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Target

TARGET K

Specifies a target. CUSUM statistics are cumulative deviations from this target, or nominal, specification.

Parameters

SIGMA K...K

SIGMA C

For variables control charts, the process is assumed to produce data from a stable population that often follows a normal distribution. SIGMA can be used to specify values for the standard deviation. The values you enter may be either actual known process parameters, or they may be estimates obtained from past data.

If you use a column, it can contain different values for each historical group. If you enter one constant, all historical stages use that value; otherwise, enter a value for each historical group.

If you do not specify σ, then Minitab estimates it from the data.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

POOLED

Specifies the pooled standard deviation to estimate σ when subgroup size > 1. The pooled standard deviation is the most efficient method of estimating standard deviation when you can assume constant variation across subgroups.

RBAR

Specifies the average of the subgroup ranges to estimate σ when subgroup size > 1. This method is the one that appears in many books and manuals that describe control charts. Subgroup size must be < 100 to use this subcommand.

SBAR

Specifies the average of the subgroup standard deviations to estimate σ when subgroup size > 1.

AMR

Specifies the average moving range to estimate σ when subgroup size = 1.

MMR

Specifies the median moving range to estimate σ when subgroup size = 1.

SRMSSD

Specifies the square root of half the mean of the squared successive differences to estimate σ when subgroup size = 1.

SUBGROUP K

If the subgroup sizes are not equal, then the control limits will not be straight lines, but will vary with the subgroup size. The center line of charts for ranges and standard deviations also varies with the subgroup size. If the sizes do not vary much, you may want to force the control limits to be straight lines. SUBGROUP allows you to do this. For example, SUBGROUP 6 says to calculate the control limits and center line as if all subgroups were of size 6. The data that are actually plotted are not affected by this subcommand.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations. The default is to use biasing constants.

UNBIASED

Uses unbiased constants in the calculation of standard deviations. Use this subcommand only in conjunction with methods of estimating σ involving the standard deviation (pooled standard deviation, SBAR, SRMSSD).

RSPAN K

Specifies the length of the moving range. By default, a span of two is used because consecutive values have the greatest chance of being alike. The span must be < 100.

Plan/Type

PLAN K K

Specifies a CUSUM plan, which is defined by the parameters h and k. The first value entered is h, the second is k. Both h and k must be greater than 0. The default values are 4 and 0.5.

VMASK K

Specifies the subgroup number to center the V-mask on. If you use 0 or do not specify K, the VMASK is placed on the last subgroup.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX C

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL
If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K
Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C
Stores means, one row for each historical stage.

VARIATION C
Stores standard deviations, one row for each historical stage.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores center line values, one row for each plotted point.

CONLIMITS C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C

Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

Multivariate Control Charts

TSGV: Session command for creating a Tsquared-generalized variance chart

TSGV C...C *E*

Draws a Tsquared chart and a generalized variance chart in the same graph, with subgroup size = E. Use the Tsquared-generalized variance chart to examine whether the process mean vector and process variability for two or more correlated variables are in control.

You must have at least two columns of correlated data. For example, you can create a Tsquared-generalized variance chart to investigate how the tensile strength and diameter of a fiber jointly affect the quality of a fabric.

If you have correlated variables, the use of separate control charts is misleading because the variables jointly affect the process. If you use separate univariate control charts in a multivariate situation, type I error and the probability of a point correctly plotting in control are not equal to their expected values. The distortion of these values increases with the number of measurement variables.

Multivariate control charting has several advantages over creating multiple univariate charts:

- The actual control region of the related variables is represented (elliptical for bivariate case).
- You can maintain a specific type I error.
- A single control limit determines whether the process is in control.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C...C

MU specifies the vector of historical means, one column for each column of data on the main command line. Any one column may have multiple values (one for each historical group) or a single value used for all historical stages. If using constants, there must the same number of Ks as columns of data on the main command line.

SIGMA *M M*

SIGMA specifies the historical covariance matrix, M...M, which must be a positive definite square matrix with the same number of rows and columns as columns of data on the main command. The matrix's diagonal elements must be positive. Element ij must equal element ji.

If only one M is entered, then all historical stages use that pair. If you want each historical stage to use its own covariance matrix, then enter the same number of matrices as historical stages.

NUMBER K...K

Use NUMBER to specify the number of samples used to estimate covariance.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

Confidence limits

SLIMITS K...K

SLIMITS C
Draws additional confidence limits at the specified values. The arguments for SLIMITS are positive numbers.

LOLOWER K
Sets bound on lower confidence limit for Tsquared chart.

LOUPPER K
Sets bound on upper confidence limit for Tsquared chart.

VALOWER K
Sets bound on lower confidence limit for generalized variance chart.

VAUPPER K
Sets bound on upper confidence limit for generalized variance chart.

Stages

HPROCESS C
HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C
If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]
The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores covariances, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

DECOMP C
Stores decomposed Tsquared values for out-of-control points.

TSQUARED: Session command for creating a Tsquared chart

TSQUARED C...C *E*
Creates a Tsquared chart, with subgroup size = E. Use the Tsquared chart to examine if the process mean vector for two or more correlated variables is in control.

You must have at least two columns of correlated data. For example, you can create a Tsquared chart to investigate the percentage of impurities and the concentration strength in a chemical process.

If you have correlated variables, the use of separate control charts is misleading because the variables jointly affect the process. If you use separate univariate control charts in a multivariate situation, type I error and the probability of a point correctly plotting in control are not equal to their expected values. The distortion of these values increases with the number of measurement variables.

- The actual control region of the related variables is represented (elliptical for bivariate case).
- You can maintain a specific type I error.
- A single control limit determines whether the process is in control.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

MU K...K

MU C...C

MU specifies the vector of historical means, one column for each column of data on the main command line. Any one column can have multiple values (one for each historical group) or a single value used for all historical stages. If using constants, there must the same number of Ks as columns of data on the main command line.

SIGMA K

SIGMA specifies the historical covariance matrix, M...M, which must be a positive definite square matrix with the same number of rows and columns as columns of data on the main command. The matrix's diagonal elements must be positive. Element ij must equal element ji.

If only one M is entered, then all historical stages use that pair. If you want each historical group to use its own covariance matrix, then enter the same number of matrices as historical stages.

NUMBER K...K

Use NUMBER to specify the number of samples used to estimate covariance.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using EXCLUDE 1), then decide to omit observation 2 (using OMIT 2). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using EXCLUDE 1), then decide to omit observation 2 (using OMIT 2). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

Confidence limits

SLIMITS K...K

SLIMITS C

Draws additional confidence limits at the specified values. The arguments for SLIMITS are positive numbers.

UPPER K

Sets a bound on the upper confidence limit.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores covariances, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C

Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

SAMPSIZE C
Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C
Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

DECOMP C
Stores decomposed Tsquared values for out-of-control points.

GENVAR: Session command for creating a generalized variance chart

GENVAR C...C *E*
Draws a generalized variance chart with subgroup size = E. Use the generalized variance chart to examine if the process variability for two or more correlated variables is in control.

You must have at least two columns of correlated data. For example, you can create a Tsquared-generalized variance chart to investigate how the tensile strength and diameter of a fiber jointly affect the quality of a fabric.

If you have correlated variables, the use of separate control charts is misleading because the variables jointly affect the process. If you use separate univariate control charts in a multivariate situation, type I error and the probability of a point correctly plotting in control are not equal to their expected values. The distortion of these values increases with the number of measurement variables.

Multivariate control charting has several advantages over creating multiple univariate charts:

- The actual control region of the related variables is represented (elliptical for bivariate case).
- You can maintain a specific type I error.
- A single control limit determines whether the process is in control.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

SIGMA *M...M*

SIGMA specifies the historical covariance matrix, M...M, which must be a positive definite square matrix with the same number of rows and columns as columns of data on the main command. The matrix's diagonal elements must be positive. Element ij must equal element ji.

If only one M is entered, then all historical stages use that pair. If you want each historical stage to use its own covariance matrix, then enter the same number of matrices as historical stages.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

Confidence limits

SLIMITS K...K

SLIMITS C
Draws additional confidence limits at the specified values. The arguments for SLIMITS are positive numbers.

LOWER K
Sets a bound on the lower confidence limit.

UPPER K
Sets a bound on the upper confidence limit.

Stages

HPROCESS C
HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C
If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]
The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless $\lambda = 0$, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

VARIATION C

Stores covariances, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C

Stores control limit values, one row for each plotted point.

STAGES C

Stores stage indicators, one row for each plotted point.

SAMPSIZE C

Stores subgroup sizes, one row for each plotted point.

TRESULTS C...C

Stores test results, one row for each plotted point. Minitab assigns each plotted point one of three values: 1 (beyond upper control limit), −1 (beyond lower control limit), 0 (in control).

MEWMA: Session command for creating a multivariate EWMA chart

MEWMA C...C *E*

Draws a multivariate EWMA chart, the multivariate equivalent to the EWMACHART on page 583. Use the multivariate EWMA chart to examine small shifts in the mean vector of two or more correlated variables.

You must have at least two columns of correlated data. For example, you can create a Tsquared-generalized variance chart to investigate how the tensile strength and diameter of a fiber jointly affect the quality of a fabric.

If you have correlated variables, the use of separate control charts is misleading because the variables jointly affect the process. If you use separate univariate control charts in a multivariate situation, type I error and the probability of a point correctly plotting in control are not equal to their expected values. The distortion of these values increases with the number of measurement variables.

Multivariate control charting has several advantages over creating multiple univariate charts:

- The actual control region of the related variables is represented (elliptical for bivariate case).
- You can maintain a specific type I error.
- A single control limit determines whether the process is in control.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

ARL and weight

`ARL K`

Sets the average run length (ARL) of K to determine the upper control limit. K must be a positive number. The default ARL is 200. The minimum allowable ARL is 1.

`WEIGHT K`

Specifies the weight. K must be between 0 and 1. The default is 0.1.

Parameters

MU C...C

MU K...K

MU specifies the vector of historical means, one column for each column of data on the main command line. Any one column may have multiple values (one for each historical group) or a single value used for all historical stages. If using constants, there must the same number of K's as columns of data on the main command line.

SIGMA *M...M*

SIGMA specifies the historical covariance matrix, M...M, which must be a positive definite square matrix with the same number of rows and columns as columns of data on the main command. The matrix's diagonal elements must be positive. Element ij must equal element ji.

If only one M is entered, then all historical stages use that pair. If you want each historical group to use its own covariance matrix, then enter the same number of matrices as historical stages.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

Confidence limits

UPPER K

Sets a bound on the upper confidence limit.

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Box-Cox transformation

BOXCOX [K...K]

BOXCOX [C]

The Box-Cox transformation can be used to correct both non-normality in process data and subgroup process variation being related to the subgroup mean. Under most conditions, it is not necessary to correct for non-normality unless the data are highly skewed.

Minitab provides two Box-Cox transformations: a stand-alone command, and a transformation option provided with each control chart. You can use these procedures in tandem. First, use the stand-alone command as an exploratory tool to help you determine the best lambda value for the transformation. Then, when you enter the control chart command, use the transformation option to transform the data at the same time you draw the chart.

You can use the Box-Cox power transformation when your data are very skewed or where the within-subgroup variation is unstable. The transformation takes the original data to the power λ, unless λ = 0, in which case the transformation takes the natural log.

You can use BOXCOX only with positive data.

You can use the common transformations natural log (λ = 0) and square root (λ = 0.5). You can also choose any value between −5 and 5 for λ or choose to find a single optimal λ for each or all historical stages. In most cases, you should not choose a λ outside the range of −2 and 2.

When you use this transformation, Minitab does not accept any values entered for MU or SIGMA.

When you use the Box-Cox transformation, the control chart is of the transformed data. The process parameters (mean and standard deviation) are also calculated using the transformed data.

ALL

If you use the ALL argument, Minitab uses data from all historical stages to find a single lambda. If you do not specify any arguments, then Minitab finds the optimal value of lambda.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores means, one row for each historical stage.

VARIATION C

Stores covariances, one row for each historical stage.

PPOINTS C

Stores plotted points, one row for each subgroup.

CONLIMITS C

Stores control limit values, one row for each plotted point.

STAGES C

Stores stage indicators, one row for each plotted point.

SAMPSIZE C

Stores subgroup sizes, one row for each plotted point.

TRESULTS C

Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Rare Event Charts

GCHART: Session command for creating a G chart

GCHART C...C

Use a G chart to monitor the number of opportunities or, in many cases, the number of days between rare events, such as infections or surgical complications.

When you monitor rare events with a traditional attribute chart, such as a P or a U chart, you need a large amount of data to establish accurate control limits. Therefore, collecting enough rare event data to detect an adverse change in the frequency of events may take months or even years. You do not need to collect large amounts of data to use a G chart. The G chart is based on the geometric distribution. The G chart plots the number of opportunities between events so you can easily detect when events occur more or less often than expected. For more information, go to How to enter data for GCHART on page 1242.

In some situations, the actual number of opportunities between events is not known. Instead, the dates of the events are recorded. When you use the G chart in these situations, you should verify that the number of possible opportunities per day is fairly constant.

If you know the date and specific time of day for each event, use TCHART on page 627 to monitor time between rare events.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

P K...K

P C

You can enter historical data for estimating p. For example, if you have a known process parameter or an estimate obtained from past data, enter the value(s). If you do not specify a value for p, Minitab estimates it from the data.

By default, Minitab estimates the chance of an event occurring in any particular day (or opportunity), p, from the overall sample proportion. Then, Minitab calculates the center line and control limits using this value.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

P limits

PLIMITS K...K

PLIMITS C

Each argument in PLIMITS defines a single line to be drawn at the specified percentile. Values entered should be percents. To get lines above and below the center line, specify matching percentiles above and below. For example, in `PLIMITS 0.1 2.5 10 90 97.55 99.99`, the first 3 percentiles are below the center line, the last 3 are above.

LOWER K

Sets lower control limit bound. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB will be drawn at the lower bound instead.

UPPER K

Sets upper control limit bound. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB will be drawn at the upper bound instead.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

BENNEYAN K

Use K=1 to perform the Benneyan test. Use K=0 to not perform the Benneyan test.

The Benneyan test counts the number of consecutive plotted points equal to 0. When a point on a G chart fails the Benneyan test, the point is marked with B. The number of points that are required to signal the Benneyan Test is a function of the desired false alarm rate and the event probability. The false alarm rate is based on the probability associated with the test 1 argument for a chart based on the normal distribution, such as an XBar chart. The default test 1 argument is 3.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]

Splits series by historical stages or by size.

NSESSION

Suppresses display of test results in the table under the graph.

INLABELS

Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS

Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

LOCATION C

Stores estimates of the proportions, one row for each historical stage.

PPOINTS C C

Stores plotted points, one row for each subgroup.

CENLINE C C

Stores center line values, one row for each plotted point.

CONLIMITS C C C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators, one row for each plotted point.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

TCHART: Session command for creating a T chart

TCHART C...C
Use a T chart to monitor the time between rare events, such as infections or surgical complications.

When you monitor rare events with a traditional variables control chart, you need a large amount of data to establish accurate control limits. Therefore, collecting enough rare event data to respond to an adverse change in the frequency of events may take months or even years. You do not need to collect large amounts of data to use a T chart. The T chart is based on the Weibull distribution or the exponential distribution, and the chart plots the time between events so you can easily detect when events occur more or less frequently than usual. For details, go to How to enter data for TCHART on page 1244.

If you know the date but not the specific time for each event, use GCHART on page 622 to monitor the number of complete days between rare events.

UNITS *"text"*
Defines the units for numeric time-between data ("Hours" or "Days").

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

STAMP: Session subcommand for specifying columns that contain time values for observations on page 1156

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CLIMITS: Session command for specifying attributes for control limit lines on page 1079

CLINE: Session command for specifying attributes for a center line on page 1081

CONNECT: Session subcommand for connecting points with lines on page 1085

HLINE: Session subcommand for specifying attributes for historical stage lines on page 1105

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Parameters

You can base the chart on either a Weibull distribution or an exponential distribution. You can also specify historical values for the distribution parameters if, for example, you have known process parameters or estimates that were obtained from past data. The following table shows how to specify the distribution and parameters.

Distribution	Command language
A Weibull distribution and estimate the shape and scale parameters from the data	None
A Weibull distribution with shape = X and scale = Y	`WSHAPE X  1` `WSCALE Y`
An exponential distribution, when you want to estimate the scale parameter from the data	`WSHAPE 1`
An exponential distribution with scale = X	`WSHAPE 1` `WSCALE X`

If your chart includes stages, and you enter one value for a parameter, then all stages use that same parameter value. You can also enter a separate parameter value for each stage.

Historical values between 0.5 and 2 are most practical for the shape of the Weibull distribution. Your historical scale value for the Weibull distribution should be slightly greater than the mean of the data. For an exponential distribution, the scale parameter is the mean.

WSHAPE K...K

WSHAPE C

Specifies the shape parameter.

WSCALE K...K

WSCALE C

Specifies the scale parameter.

Estimate

OMIT K...K

OMIT C

Use OMIT to specify subgroups to omit from the calculations for the estimates of μ and σ.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `OMIT 10:30 32` specifies to omit subgroups 10 to 30 and subgroup 32. Or suppose you have 50 subgroups, but the 24th is an outlier with a corrected assignable cause (that is, you do not want to use the 24th in parameter estimation). You should use `OMIT 24`. OMIT with a column is useful in macros.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with OMIT.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

ESTIMATE K...K

ESTIMATE C

Use ESTIMATE to list the subgroups to be used.

You cannot use OMIT and ESTIMATE in the same command.

List the subgroups, or a column that contains the subgroup numbers that you want to omit. Use a colon to indicate a range of values or columns. For example, `ESTIMATE 10:30 32` specifies to use subgroups 10 to 30 and subgroup 32.

When you use the control chart dialogs to create the chart, you can use only constants (not columns) with ESTIMATE.

Important If you use OMIT or ESTIMATE with INCLUDE or EXCLUDE, the estimate or used samples apply to the entire data set, not just the subsetted data. Suppose you exclude observation 1 from your analysis (using `EXCLUDE 1`), then decide to omit observation 2 (using `OMIT 2`). Because Minitab considers an omitted observation or sample (observation 1) to still be part of the entire data set, you must enter observation 2 as the observation to omit.

P limits

PLIMITS K...K

PLIMITS C

Each argument in PLIMITS defines a single line to be drawn at the specified percentile. Values entered should be percents. To get lines above and below the center line, specify matching percentiles above and below. For example: `PLIMITS 0.1 2.5 10 90 97.55 99.99`. The first 3 percentiles are below the center line, and the last 3 are above.

LOWER K

Sets the lower control limit bound. K must be a positive number. If the calculated lower control limit is less than the lower bound, a horizontal line labeled LB is drawn at the lower bound instead.

UPPER K

Sets the upper control limit bound. K must be a positive number. If the calculated upper control limit is greater than the upper bound, a horizontal line labeled UB is drawn at the upper bound instead.

Tests

TEST K...K

TEST C

Performs one or more of the eight tests for special causes. Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

To adjust the sensitivity of the tests, use the DEFTEST subcommand.

To use TEST, list the numbers of the tests you want to use, or a column which contains the numbers of the tests you want to use. For example, `TEST 1 3` specifies perform tests 1 and 3. `TEST 1:8` specifies do all eight tests. TEST with a column is useful when writing macros. Test 0 will perform no tests and override any set control chart options.

When a point fails a test, it is marked with the test number on the chart. If a point fails more than one test, the number of the first test in your list is the number displayed on the chart.

Minitab will perform tests when subgroup sizes are unequal.

DEFTEST

Defines the sensitivity of the tests for special causes used with quality control charts.

The test definitions only apply to this chart.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K: Test for special causes	Description of the test	Second K: Sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line (Minitab also draws the control limits at K standard deviations from the center line.)	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Stages

HPROCESS C

HPROCESS lets the user define historical stages in the process data. The control limits and center line are then estimated independently within each stage. Stages are much like groups in other Minitab graph commands. Thus, the HPROCESS column is like a group variable in other graph commands, with one exception—when the value in the column changes, a new stage starts. This works the same way the subgroup column works in the control chart commands.

HPROCESS is often done for comparative purposes. For example, suppose that a process improvement is made at some point in time. This usually results in a reduction of the process variation. It may be of interest to draw a control chart containing data prior to the improvement and data after the improvement so that one can directly compare the two to see the change in process variation.

VALUES K...K

VALUES C

If you use the optional VALUES sub-subcommand, then Minitab uses the values listed in that column or list of constants to estimate new parameters.

Display

DISPLAY K

Displays only the final K subgroups in the control chart.

SPLIT [K]
Splits series by historical stages or by size.

NSESSION
Suppresses display of test results in the table under the graph.

INLABELS
Displays the center lines and control limits for each stage, inside the data region. You can specify the exact position of particular control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

OUTLABELS
Displays the center lines and control limits for the final stage, outside the data region. You can specify the exact position of the control limits or center lines using the POSITION / ENDPOSITION subcommands. However, it may be easier to just move them with the pointer.

Storage

SSHAPE C
Stores the shape parameter.

SSCALE C
Stores the scale parameter.

PPOINTS C
Stores plotted points, one row for each subgroup.

CENLINE C
Stores the center line value.

CONLIMITS C C
Stores control limit values, one row for each plotted point.

STAGES C
Stores stage indicators.

TRESULTS C...C
Stores test results, one column for each test performed and one row for each plotted point. Minitab assigns each plotted point a value: 1 (failed test) or 0 (did not fail test).

Tests for Special Causes

DEFTEST: Session command for defining the sensitivity of the tests for special causes

DEFTEST K K...K K
Defines the sensitivity of the tests for special causes that are used with control charts.

DEFTEST applies to any chart which supports the TEST subcommand. The first K in each pair corresponds to the number of the test; the second K defines the test.

First K, the number of the test for special causes	Description of the test	Range of acceptable values for the second K, the sensitivity of the test	Default value of second K
1	One point more than K standard deviations from center line For standard control limits, Minitab draws the control limits at K standard deviations from the center line. When you use extended limits, Test 1 identifies a single point outside the extended limits. The value of K does not apply.	1–6	3
2	K points in a row on same side of center line	7–11	9
3	K points in a row, all increasing or all decreasing	5–8	6
4	K points in a row, alternating up and down	12–14	14
5	K out of K+1 points more than 2 standard deviations from center line (same side)	2–4	2
6	K out of K+1 points more than 1 standard deviation from center line (same side)	3–6	4
7	K points in a row within 1 standard deviation of center line (either side)	12–15	15
8	K points in a row more than 1 standard deviation from center line (either side)	6–10	8

Capability Analysis

DCAPA: Session command for performing individual distribution identification

DCAPA [C] [C]

DCAPA [C] [K]

Fits data in a column or in subgroups across rows of several columns with 14 distributions: normal, lognormal, 3-parameter lognormal, gamma, 3-parameter gamma, exponential, 2-parameter exponential, smallest extreme value, weibull, 3-parameter weibull, largest extreme value, logistic, loglogistic, and 3-parameter loglogistic; and 2 transformations: Box-Cox and Johnson. You can specify subgroups in another column or using a constant.

The report includes goodness-of-fit test and probability plot to assess whether a distribution fits your data.

The values in the column must be numeric. Columns can have missing values.

RSUB C...C

Specifies that the subgroup data for a variable are in rows of several columns. Columns must be the same length. You can use this subcommand only when no column is specified with the DCAPA. For more information, go to How to enter subgroup data on page 1246.

Distributions and transformations

Use the following subcommands to specify which distributions to fit your data with.

NORMAL
The normal distribution.

LNORMAL
The lognormal distribution.

TLNORMAL
The 3-parameter lognormal distribution.

EXPONENTIAL
The exponential distribution.

TEXPONENTIAL
The 2-parameter exponential distribution.

WEIBULL
The Weibull distribution.

TWEIBULL
The 3-parameter Weibull distribution.

LEXTREMEVALUE
The largest extreme value distribution.

SEXTREMEVALUE
The smallest extreme value distribution.

GAMMA
The gamma distribution.

TGAMMA
The 3-parameter gamma distribution.

LOGISTIC
The logistic distribution.

LLOGISTIC
The loglogistic distribution.

TLLOGISTIC
The 3-parameter loglogistic distribution.

BOXCOX
Transforms the data to follow normal distribution using Box-Cox transformation. Specify the value of lambda as K. K must be between −5 and 5. If no value is specified, Minitab estimates the optimal value of lambda from the data. Box-Cox transformation is valid only for positive data values.

The output includes goodness-of-fit test and normal probability plot for original and transformed data for comparison.

STORE C
Stores the transformed data in C, when data are in one column.

STORE C...C
Stores the data in C...C, when data are in subgroups.

JOHNSON
Transforms the data to follow normal distribution using Johnson transformation. Specify the value of the P-value as K. K must be between 0 and 1. If no value is specified, Minitab uses a default P-value of 0.01.

The output includes goodness-of-fit test and normal probability plot for original and transformed data for comparison.

STORE C
Stores the transformed data in C, when data are in one column.

STORE C...C
Stores the data in C...C, when data are in subgroups.

ALL
Use ALL when you want all 14 distributions but want to specify values for the data transformations or specify storage options. For example, if you want a specific value of lambda, say λ = 2, and you want to store the transformed data, then use:

```
ALL;
BOXCOX 2;
 STORE C.
```

Results

RDESCRIPTIVE
Displays the descriptive statistics for the data column.

RFITTESTS
RFITTESTS performs a goodness-of-fit test for your choice of distribution. The session output contains Anderson-Darling (AD), corresponding p-value (P), and likelihood ratio test p-value (LRT P).

SPVALUE C C C
SPVALUE stores the name of the distribution and corresponding p-values of the Anderson-Darling and Anderson-Darling likelihood-ratio test in C, C, and C, respectively.

RPTILE
Specifies the percentiles to be estimated. K and values in C lie between 1 and 100.

RESTIMATE
Displays the estimated distribution parameters.

Options

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

SEPARATE
Displays one probability plot per graph. The default is 4 probability plots per graph.

NOCI
Displays probability plots with no confidence intervals.

JOHNSON: Session command for applying the Johnson transformation

JOHNSON [C]
Transforms your data to follow normal distribution when feasible. Specify a column when data are in a single column. Use RSUB when data are in subgroups across rows of several columns.

The values in the column must be numeric and contain at least four distinct observations.

RSUB C...C
Specifies that the subgroup data for a variable are in rows of C...C. The values in the columns must be numeric and contain at least four distinct observations.

STORE C
Stores the transformed data in column C when the input data are in one column.

STORE C...C
Stores the transformed data in subgroups in rows of C...C. The number of columns specified must equal to that specified with RSUB.

SPARAMETERS
Stores the optimal distribution type and parameters estimates in a column. The first row in C stores distribution type, SB, SL, or SU, the second row gamma parameter estimate, the third row eta, the fourth row epsilon, and the five row lambda (when available).

Options

PVALUE K
Specifies a p-value for the transformation.

BWCAPA: Session command for performing between/within capability analysis

BWCAPA [C...C *E*]
Performs a capability analysis using within and between subgroup variations on data sets in C...C with subgroups in E. Performs a capability analysis and generates a process capability report for data from a normal distribution using both between-subgroup variation and within-subgroup variation.

If the data set is arranged in columns, specify the columns in C...C. Columns can contain numeric or text data. Specify the subgroup size in E. E can be a column (if the subgroup size varies) or a constant (if the subgroups are the same size). If your data are in subgroups across rows of several of columns, use RSUBS. For more information, go to How to enter data.

The process capability report includes the following components:

- A capability histogram that is overlaid with two normal curves. Minitab generates the two normal curves using the process mean and the overall standard deviation, and the process mean and the total standard deviation.

- A complete table of overall and total (between- and within-subgroup) capability statistics. To calculate the statistics, Minitab uses either the historical values that you specify or the values that it estimates from the data for within- and between-subgroup standard deviations.
- Capability statistics to assess whether the process is capable of meeting the specifications consistently.

RSUBS C...C [K]

Specifies that subgroup data for a variable are in rows of C...C. All columns must be the same length. If you specify only one input column with RSUBS, then the subgroup size is 1. If the number of data sets is greater than 1, then you can use K to specify the number. For more information, go to How to enter data.

USPEC C

USPEC K...K

Specifies the upper specification limit for each data set. You can specify values for a limit in a column (C) or as constants (K...K), one for each data set used in the analysis.

Note You must enter either a value for either USPEC or LSPEC.

UBOUND C

UBOUND K...K

Specifies that the upper specification limit is a boundary in C or K...K.

Values in C or K...K	Limit
0	The limit is not a boundary.
1	The limit is a boundary.

LSPEC C

LSPEC K...K

Specifies the lower specification limit for each data set. You can specify values for a limit in a column (C) or as constants (K...K), one for each data set used in the analysis.

Note You must enter either a value for either USPEC or LSPEC.

LBOUND C

LBOUND K...K

Specifies that the lower specification limit is a boundary in C or K...K.

Values in C or K...K	Limit
0	The limit is not a boundary.
1	The limit is a boundary.

Historical estimates

It is generally assumed that the data from the process follow a normal distribution. MU is the mean of this distribution. BSIGMA is the process standard deviation between batches and WSIGMA is the process standard deviation within subgroups. You can use these subcommands to enter known values for the process parameters, or enter previous estimates. If you omit these subcommands, then Minitab estimates the values from the data sets.

You can specify the values in a column or as constants, one for each data set used in the analysis.

MU C

MU K...K
Specifies historical values for the mean of the normal distribution for each data set in C or as K...K. If you don't specify a value, then Minitab estimates it from the data.

BSIGMA C

BSIGMA K...K
Specifies historical values for the between-subgroup standard deviation for each data set in C or as K...K. If you don't specify a value, then Minitab estimates it from the data.

WSIGMA C

WSIGMA K...K
Specifies historical values for the within-subgroup standard deviation for each data set in C or as K...K. If you don't specify a value, then Minitab estimates it from the data.

Options

TARGET C

TARGET K...K
Specifies a process target value in C or as K...K, one for each data set used in the analysis. Minitab then calculates Cpm, in addition to the standard capability statistics.

TOLER K
Specifies a value to calculate the capability statistics using an interval that is 6 standard deviations wide, 3 on either side of the process mean. TOLER lets you change the width of this interval. For example, `TOLER 12` says to use an interval 12 standard deviations wide, 6 on either side of the process mean. This is referred to as "6-sigma quality".

CSIXSIGMA
Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA
Specifies the AIAG/VDA capability naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE
Specifies the capability indices for process performance.

PCAPABILITY
Specifies the capability indices for long-term capability.

MPERFORMANCE
Specifies the capability indices for machine performance.

WITHIN
Performs within-subgroup capability analysis only. If you don't include this subcommand, then Minitab performs both within-subgroup capability analysis and overall capability analysis.

OVERALL
Performs overall capability analysis only. If you don't include this subcommand, then Minitab performs both within-subgroup capability analysis and overall capability analysis.

PERCENT
Uses percentages instead of the default parts per million (PPM).

PPM
Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

ZBENCH
Displays benchmark Z statistics instead of the default capability statistics.

CSTAT
Displays the capability statistics when you include ZBENCH to display benchmark Z statistics.

CONFIDENCE [K]
Specifies a confidence level in K to calculate the confidence intervals for Cp, Cpk, Pp, Ppk, and Cpm statistics. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ONECI
Specifies one-sided confidence intervals. The direction of the intervals depends on the statistic.

Lower bound:

- Cp
- Cpk
- Ppk
- Cpm
- Z.bench

Upper bound:

- Parts Per Million (PPM) < Lower Specification Limit (LSL)
- PPM > Upper Specification Limit (USL)
- PPM total
- Percent < LSL
- Percent > USL
- Percent total

TWOCI
Specifies two-sided confidence intervals.

NOCI
Specifies not to include confidence intervals in the output.

Box-Cox transformation

BOXCOX [C]

BOXCOX [K...K]

Transforms the data to follow a normal distribution using the Box-Cox transformation.

You can specify lambda values that you want Minitab to use in the transformation in C or K...K, one value for each data set used in the analysis. Values that you specify must be between –5 and 5. If you don't specify lambda values, Minitab estimates the optimal value of lambda from the data.

Sigma estimates

AMR

Estimates the within standard deviation using the average moving range with the length that you specify with RSPAN in K.

RSPAN K

Specifies the length that you want to use for the moving range. The maximum length is 100.

RBAR

Estimates within standard deviation using the average of subgroup range method.

SBAR

Estimates within standard deviation using the averages of the subgroup standard deviation method.

MMR

Estimates within standard deviation using the median of the moving ranges with the length that you specify with RSPAN in K.

RSPAN K

Specifies the length that you want to use for the moving range. The maximum length is 100.

SRMSSD

Estimates the within standard deviation using the square root of half of the mean of the squared successive differences.

POOLED

Estimates the within standard deviation using the pooled standard deviation.

BIASED

Specifies to not use unbiasing constants in the calculation of the within-subgroup standard deviations. The default is to use unbiasing constants.

UNBIASED

Specifies to use unbiasing constants in the calculation of the within-subgroup standard deviations.

OBIASED

Specifies to not use an unbiasing constant in the calculation of the overall standard deviation. The default is to use an unbiasing constant.

OUNBIASED

Specifies to use an unbiasing constant in the calculation of the overall standard deviation.

Graphs

NOCHART
Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores the name of the columns of measurements used for the analysis.

USL C
Stores the upper specification limit.

STARGET C
Stores the target of the process.

LSL C
Stores the lower specification limit.

MEAN C
Stores the process mean.

N C
Stores the sample size used for the analysis.

SDB C
Stores the between-subgroup standard deviation.

SDW C
Stores the within-subgroup standard deviation.

SDT C
Stores the between/within-subgroup standard deviation of the transformed data.

SDO C
Stores the overall standard deviation.

CP C
Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cp or Z.Bench
CAIAGVDA	Cw or Zw.Bench

CPUC C
Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCp or Z.Bench
CAIAGVDA	UBCw or Zw.Bench

CPLC C
Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCp or Z.Bench

If you specify	Minitab stores
CAIAGVDA	LBCw or Zw.Bench

CPU C

Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	CPU or Z.USL
CAIAGVDA	CwU or Zw.USL

CPL C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	CPL or Z.LSL
CAIAGVDA	CwL or Zw.LSL

CPK C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	Cpk
CAIAGVDA	Cwk

CPKU C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCpk
CAIAGVDA	UBCwk

CPKL C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCpk
CAIAGVDA	LBCwk

CCPK C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	CCpk

If you specify	Minitab stores
CAIAGVDA	CCwk

PP C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Zm.Bench

PPUC C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPm.G or Zm.Bench

PPLC C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPm.G or Zm.Bench

PPU C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPU or Z.USL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpU.Z or Zp.USL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpU.Z or CZp.USL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmU.Z or Zm.USL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpU.G or Zp.USL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpU.G or CZp.USL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmU.G or Zm.USL

PPL C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPL or Z.LSL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpL.Z or Zp.LSL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpL.Z or CZp.LSL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmL.Z or Zm.LSL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpL.G or Zp.LSL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpL.G or CZp.LSL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmL.G or Zm.LSL

PPK C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G

PPKU C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmk.G

PPKL C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmk.G

CPM C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Cpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmm.G

UCPM C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBCpm

If you specify	Minitab stores
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmm.G

LCPM C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBCpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmm.G

PPML C

Store the observed parts per million less than the lower specification limit.

PPMU C

Stores the observed parts per million greater than the upper specification limit.

PPMT C

Stores the total parts per million beyond the specification limit.

EPLO C (C C)

Stores either the expected parts per million or the expected percentage less than the lower specification limit. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUO C (C C)

Stores either the expected parts per million or the expected percentage greater than the upper specification limit. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTO C (C C)

Stores either the expected parts per million or the expected percentage outside the specification limits. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPLW C (C C)
Stores either the expected parts per million or the expected percentage less than the lower specification limit. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUW C (C C)
Stores either the expected parts per million or the expected percentage greater than the upper specification limit. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTW C (C C)
Stores either the expected parts per million or the expected percentage outside the specification limits. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

SLAMBDA C
Stores the lambda value that Minitab uses in the Box-Cox transformation.

TUSL C
Stores the upper specification limit of the Box-Cox transformed data.

TTARGET C
Stores the target value of the Box-Cox transformed data.

TLSL C
Stores the lower specification limit of the Box-Cox transformed data.

TMEA C
Stores the mean of the Box-Cox transformed data.

TSDB C
Stores the between-subgroup standard deviation of the Box-Cox transformed data.

TSDW C
Stores the within-subgroup standard deviation of the Box-Cox transformed data.

TSDT C
Stores the total standard deviation of the Box-Cox transformed data.

TSDO C
Stores the overall standard deviation of the Box-Cox transformed data.

CAPA: Session command for performing a normal capability analysis

CAPA [C...C *E*]
Performs a capability analysis based on the assumption that data come from normal population. You can also use the BOXCOX or JOHNSON subcommand to try to transform your data to follow normal distribution. If a data set is arranged in a column, specify the columns with subgroup size in E. E can be a column (if the subgroup size varies) or a constant (if the subgroups are the same size.) Columns may contain numeric or text data. If your data

are in subgroups across rows of several of columns, use RSUBS. For details, go to How to enter subgroup data on page 1246.

CAPA displays a histogram, overlaid with a normal curve generated from the process mean and standard deviation. It also displays the locations of the process mean and the target, the natural process tolerance, and the process specifications. Hence, the chart can be used to visually assess whether the data are normally distributed, whether the process is centered on the target, and whether it is capable of meeting the process specifications consistently.

RSUBS C...C [K]

Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUB, then the subgroup size is 1.

The optional K indicates the number of data sets if greater than 1. For details, go to How to enter subgroup data on page 1246.

USPEC C

USPEC K...K

Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

UBOUND C

UBOUND K...K

Specifies that the upper specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

LSPEC C

LSPEC K...K

Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LBOUND C

LBOUND K...K

Specifies that the lower specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

Historical estimates

MU C

MU K...K

Specifies the mean of the population distribution (denoted as μ or mu). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data. You can either specify one value applicable to all the data sets or multiple values, one value for each data set used in the analysis.

You cannot use MU with a transformation.

If you do not specify μ, then Minitab estimates it from the data.

SIGMA C

SIGMA K...K

Specifies the standard deviation of the population distribution (denoted as σ or sigma). The value that you enter can be either an actual known process parameter, or an estimate obtained from past data. You can either specify one value applicable to all the data sets or multiple values, one value for each data set used in the analysis.

You cannot use SIGMA with a transformation.

If you do not specify σ, then Minitab estimates it from the data.

Options

TARGET C

TARGET K...K

Specifies a process target value. Minitab then calculates Cpm, in addition to the standard capability statistics.

You can specify the values in a column or as constants, one value for each data set used in the analysis.

TOLER K

Specifies a value to calculate the capability statistics using an interval other than 6 standard deviations wide, 3 on either side of the process mean. For example, `TOLER 12` says to use an interval 12 standard deviations wide, 6 on either side of the process mean.

You can specify the values in a column or as constants, one value for each data set used in the analysis.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX
Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX
Calculates the capability indices using the ISO method, also called the geometric method.

WITHIN
Use WITHIN to perform a within capability analysis only. This subcommand cannot be used with a Johnson transformation.

The default is to conduct both within and overall capability analyses.

OVERALL
Use OVERALL to perform an overall capability analysis only. If a Johnson transformation is used, only an overall analysis can be done so this subcommand is redundant.

The default is to conduct both within and overall capability analyses.

PERCENT
Replaces the default parts per million (PPM) calculations with percentages.

PPM
Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

ZBENCH
Displays benchmark Z statistics rather than the default capability statistics.

CSTAT
Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

CONFIDENCE K
Specifies a confidence level in K. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ONECI
Specifies one-sided confidence intervals. The direction of the intervals depends on the statistic.

Lower bound:

- Cp
- Cpk
- Ppk
- Cpm
- Z.bench

Upper bound:

- Parts Per Million (PPM) < Lower Specification Limit (LSL)
- PPM > Upper Specification Limit (USL)
- PPM total
- Percent < LSL
- Percent > USL

- Percent total

TWOCI
Specifies two-sided confidence intervals.

NOCI
Specifies not to include confidence intervals in the output.

Box-Cox transformation

BOXCOX [C]

BOXCOX [K...K]
Transforms the data to follow a normal distribution using Box-Cox transformation. You can also specify a value for lambda to be used in the transformation. If no value is specified, Minitab estimates the optimal value of lambda from the data.

You can specify the values for lambda in a column (C) or as constants (K), one value for each data set used in the analysis. K and values in C must lie between -5 and 5.

You cannot use BOXCOX with a Johnson transformation.

Johnson transformation

JOHNSON
Selects a distribution function from Johnson system to transform each data set to follow normal distribution.

You cannot use JOHNSON with a Box-Cox transformation.

PVALUE K...K

PVALUE C
Use the optional PVALUE to change the default p-value of 0.10 to select the optimal transformation. You can specify p-values in a column or as constants, one value for each data set used in the analysis.

Standard deviation estimates

These subcommands provide options for estimating the within process standard deviation based on the following methods.

Except for OBIASED and OUNBIASED, you cannot use the following subcommands with a Johnson transformation.

AMR
Average moving range.

RSPAN K
Use RSPAN to change the length of the moving range. The maximum length allowed is 100.

RBAR
Averages of the subgroup ranges. The maximum subgroup size allowed is 100.

SBAR
Averages of the subgroup standard deviations.

MMR
Median of the moving range.

RSPAN K
Use RSPAN to change the length of the moving range. The maximum length allowed is 100.

SRMSSD
Square root of half of the mean of the squared successive differences.

POOLED
Pooled standard deviation.

BIASED
Use BIASED to suppress the use of unbiasing constants in the calculation of the within-subgroup standard deviations. The default is to use unbiasing constants.

UNBIASED
Use an unbiasing constant in the calculation of the within subgroup standard deviations.

OBIASED
Use OBIASED to suppress the use of an unbiasing constant in the calculation of the overall standard deviation. The default is to use unbiasing constants.

OUNBIASED
Use an unbiasing constant in the calculation of the overall standard deviation.

Graphs

NOCHART
Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores the name of the data columns used in the analysis.

USL C
Stores the upper specification limit.

STARGET C
Stores the process target.

LSL C
Stores the lower specification limit.

MEAN C
Stores the mean of the data.

N C
Stores the number of sample observations.

SDT C
Stores the within-subgroup standard deviation (produces missing values if used with the Johnson transformation).

SDO C

Stores the overall standard deviation.

CP C

Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cp or Z.Bench
CAIAGVDA	Cw or Zw.Bench

CPUC C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCp or Z.Bench
CAIAGVDA	UBCw or Zw.Bench

CPLC C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCp or Z.Bench
CAIAGVDA	LBCw or Zw.Bench

CPU C

Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	CPU or Z.USL
CAIAGVDA	CwU or Zw.USL

CPL C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	CPL or Z.LSL
CAIAGVDA	CwL or Zw.LSL

CPK C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	Cpk
CAIAGVDA	Cwk

CPKU C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCpk
CAIAGVDA	UBCwk

CPKL C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCpk
CAIAGVDA	LBCwk

CCPK C

Stores the following statistics, depending on the naming convention method.

If you specify	Minitab stores
CSIXSIGMA	CCpk
CAIAGVDA	CCwk

PP C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Zm.Bench

PPUC C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPm.Z or Zm.Bench

If you specify	Minitab stores
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPm.G or Zm.Bench

PPLC C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPm.G or Zm.Bench

PPU C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPU or Z.USL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpU.Z or Zp.USL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpU.Z or CZp.USL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmU.Z or Zm.USL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpU.G or Zp.USL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpU.G or CZp.USL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmU.G or Zm.USL

PPL C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPL or Z.LSL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpL.Z or Zp.LSL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpL.Z or CZp.LSL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmL.Z or Zm.LSL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpL.G or Zp.LSL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpL.G or CZp.LSL

If you specify	Minitab stores
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmL.G or Zm.LSL

PPK C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G

PPKU C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmk.G

PPKL C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmk.G

CPM C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Cpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmm.G

UCPM C

Stores the upper limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBCpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmm.G

LCPM C

Stores the lower limit of a confidence interval for the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBCpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmm.G

PPML C

Stores the observed parts per million less than the lower specification limit.

PPMU C
Stores the observed parts per million greater than the upper specification limit.

PPMT C
Stores the total parts per million beyond the specification limit.

EPLO C (C C)
Stores either the expected parts per million or the expected percentage less than the lower specification limit. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUO C (C C)
Stores either the expected parts per million or the expected percentage greater than the upper specification limit. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTO C (C C)
Stores either the expected parts per million or the expected percentage outside the specification limits. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPLW C (C C)
Stores either the expected parts per million or the expected percentage less than the lower specification limit. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUW C (C C)
Stores either the expected parts per million or the expected percentage greater than the upper specification limit. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTW C (C C)
Stores either the expected parts per million or the expected percentage outside the specification limits. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

JTYPE C
Stores the optimal distribution function type selected for Johnson transformation.

JGAMMA C
Stores the gamma parameter of a distribution function in Johnson transformation.

JETA C
Stores the eta parameter of a distribution function in Johnson transformation.

JLAMBDA C
Stores the lambda parameter of a distribution function in Johnson transformation.

JEPSILON C
Stores the epsilon parameter of a distribution function in Johnson transformation.

SLAMBDA C
Stores the value of lambda used in Box-Cox transformation.

TUSL C
Stores the Box-Cox or Johnson transformed value of the upper specification limit.

TTARGET C
Stores the Box-Cox or Johnson transformed target value.

TLSL C
Stores the Box-Cox or Johnson transformed value of the lower specification limit.

TMEA C
Stores the mean of the Box-Cox or Johnson transformed data.

TSDT C
Stores the total standard deviation of the Box-Cox or Johnson transformed data.

TSDO C
Stores the overall standard deviation of the Box-Cox or Johnson transformed data.

NNCAPA: Session command for performing a nonnormal capability analysis

NNCAPA C...C
Generates a process capability report when your data do not follow normal distribution. The report is based on your choice of a distribution.

The report includes a capability histogram overlaid with a distribution curve, and a table of capability statistics. The distribution curve and the capability statistics are based on the distribution parameter. The report also includes statistics of the process data, such as the mean, parameters, target (if you enter one), and process specifications; and the observed and expected overall performance. So the report can be used to visually assess the distribution of the process relative to the target, whether the data follows a specified distribution, and whether the process is capable of meeting the specifications consistently.

When using the nonnormal model, Minitab calculates the overall capability statistics, Pp, Ppk, PPU, and PPL. If no historical values for the parameters are specified, Minitab estimates the parameters from the data using maximum likelihood method.

RSUBS C...C [K]
Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUB, then the subgroup size is 1.

The optional K indicates the number of data sets if greater than 1. For details, go to How to enter subgroup data on page 1246.

USPEC C

USPEC K...K
Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

UBOUND C

UBOUND K...K
Specifies that the upper specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

LSPEC C

LSPEC K...K
Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LBOUND C

LBOUND K...K
Specifies that the lower specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

Distributions

Use the following subcommands to specify which distributions to fit your data with.

EXPONENTIAL
The exponential distribution.

TEXPONENTIAL
The 2-parameter exponential distribution.

LNORMAL
The lognormal distribution.

TLNORMAL
The 3-parameter lognormal distribution.

FNORMAL
The folded normal distribution.

WEIBULL
The Weibull distribution.

TWEIBULL
The 3-parameter Weibull distribution.

LEXTREMEVALUE
The largest extreme value distribution.

SEXTREMEVALUE
The smallest extreme value distribution.

GAMMA
The gamma distribution.

TGAMMA
The 3-parameter gamma distribution.

LOGISTIC
The logistic distribution.

LLOGISTIC
The loglogistic distribution.

TLLOGISTIC
The 3-parameter loglogistic distribution.

Parameters

SETS K...K

SETS C
Sets the shape parameter for the gamma and the Weibull distributions and the scale parameter for all other distributions. Minitab calculates maximum likelihood estimates from the data for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

SETTHRESHOLD K...K

SETTHRESHOLD C
Sets the threshold parameter for the 2-parameter exponential and any 3-parameter distributions. Minitab calculates maximum likelihood estimates from the data set for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

HPARAMETERS K...K

HPARAMETERS C

HPARAMETERS C...C
Sets all the parameters in a distribution. If all data sets share the same parameter values, enter constants or a column. The number of constants or values in a column must equal the number of parameter in a distribution. If the data set has different parameter values, enter columns equal to the number of data sets in your analysis. The number of entries in each column must equal the number of parameters in a distribution.

You must enter parameter values in the following order: location, shape, scale, and threshold. For example, for 3-parameter loglogistic distribution, enter values of the parameters in the following order: location, scale, and threshold.

Capability calculation method

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX
Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX
Calculates the capability indices using the ISO method, also called the geometric method.

Options

TARGET K...K

TARGET C

Specifies a process target value. Minitab then calculates Cpm, in addition to the standard capability statistics. The target and Cpm values are displayed in the table below the graph.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

TOLER K...K

TOLER C

Specifies a value to calculate the capability statistics using an interval other than 6 standard deviations wide, 3 on either side of the process mean. For example, `TOLER 12` says to use an interval 12 standard deviations wide, 6 on either side of the process mean. This is referred to as "6-sigma quality."

You can specify the values in a column or as constants, one value for each data set used in the analysis.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

ZBENCH

Displays benchmark Z statistics rather than the default capability statistics.

CSTAT

Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

PERCENT

Replaces the default parts per million (PPM) calculations with percentages.

PPM
Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

Graphs

NOCHART
Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores the name of the data columns used in the analysis.

LSL C
Stores the lower specification limit.

USL C
Stores the upper specification limit.

STARGET C
Stores the process target.

MEAN C
Stores the mean of the data.

N C
Stores the number of sample observations.

SSHAPE C
Stores the shape parameter of a distribution.

SSCALE C
Stores the scale parameter of a distribution.

SLOCATION C
Stores the location parameter of a distribution.

STHRESHOLD C
Stores the threshold parameter of a distribution.

PP C
Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Zm.Bench

PPU C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPU or Z.USL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpU.Z or Zp.USL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpU.Z or CZp.USL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmU.Z or Zm.USL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpU.G or Zp.USL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpU.G or CZp.USL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmU.G or Zm.USL

PPL C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPL or Z.LSL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpL.Z or Zp.LSL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpL.Z or CZp.LSL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmL.Z or Zm.LSL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpL.G or Zp.LSL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpL.G or CZp.LSL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmL.G or Zm.LSL

PPK C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G

PPML C

Stores the observed parts per million less than the lower specification limit.

PPMU C

Stores the observed parts per million greater than the upper specification limit.

PPMT C
Stores the total parts per million beyond the specification limit.

EPLO C
Stores the expected overall parts per million less than the lower specification limit.

EPUO C
Stores the expected overall parts per million greater than the upper specification limit.

EPTO C
Stores the expected overall parts per million beyond the specification limits.

NPCAPA: Session command for performing a nonparametric capability analysis

NPCAPA [C]
Generates a process capability report that makes no assumptions about the distribution of the data. Use a single column with the main command to enter all of the data in a single column. Enter the main command without an argument and use the RSUBS subcommand to enter data from multiple columns.

The report includes a capability histogram and a table of capability statistics. The report also includes statistics of the process data, such as the mean, median, target (if you enter one), process specifications, and the observed and expected overall performance. Use the histogram to visually assess whether the process is capable of meeting the specifications consistently.

When using the nonparametric analysis, Minitab Statistical Software calculates the overall capability statistics, Cnp, Cnpl, Cnpu, and Cnpk.

RSUBS C...C
Specifies that the data for the analysis are in multiple columns. The results are the same as if the data are all in a single column.

USPEC K
Specifies the upper specification limit. The analysis requires at least 1 specification limit.

LSPEC K
Specifies the lower specification limit. The analysis requires at least 1 specification limit.

Options

TARGET K
Specifies a process target value. Minitab Statistical Software then includes a reference line for the target on the histogram. The target is also in the table of process data.

TOLER K
Specifies a tolerance value to use to calculate the capability statistics. The value that you specify is a number of standard deviations around the mean in a normal distribution. The value defines percentiles. The analysis uses the corresponding empirical percentiles from the data to calculate the tolerance. For example, a value of 6 defines percentiles that are 3 standard deviations to either side of the mean for a normal distribution. In a normal

distribution, 3 standard deviations below the mean is the 0.15 percentile and 3 standard deviations above the mean is the 99.85 percentile. The analysis uses 0.15 empirical percentile and the 99.85 empirical percentile to calculate the capability statistics.

The default value if you do not specify the subcommand depends on the value in **File** > **Options** > **Control Charts and Quality Tools** > **Capability Analysis**. The value with the default installation of Minitab Statistical Software is 6.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. CSIXSIGMA matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use CAIAGVDA with CSIXSIGMA. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use CAIAGVDA with CSIXSIGMA. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

PERCENT

Displays the observed performance statistics in percentages. Specify only 1 of PERCENT and PPM.

PPM

Displays the observed performance statistics in parts per million (PPM). Specify only 1 of PERCENT and PPM.

ONECI

Specifies one-sided, upper confidence intervals for the observed performance statistics.

TWOCI

Specifies two-sided confidence intervals for the observed performance statistics.

NOCI

Specifies not to include confidence intervals in the output.

CONFIDENCE K

Specifies a confidence level in K. For example, for a 90% confidence level, enter CONFIDENCE 90. The default value of K is 95.

Graphs

NOCHART
Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify displays in the navigator, on the output tab, and at the top of the output above the graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores the name of the data columns used in the analysis.

LSL C
Stores the lower specification limit that you specify with LSPEC.

USL C
Stores the upper specification limit that you specify with USPEC.

STARGET C
Stores the process target that you specify with TARGET.

MEAN C
Stores the mean of the data.

MEDIAN C
Stores the median of the data.

N C
Stores the number of sample observations.

SDO C
Stores the sample standard deviation.

CNP C
Stores the following statistics, depending on the naming convention and type of process stage.

If you specify	Minitab stores
CSIXSIGMA	Cnp
CAIAGVDA and PPERFORMANCE	Pnp
CAIAGVDA and PCAPABILITY	Cnp
CAIAGVDA and MPERFORMANCE	Pnm

CNPU C
Stores the following statistics, depending on the naming convention and type of process stage.

If you specify	Minitab stores
CSIXSIGMA	Cnpu
CAIAGVDA and PPERFORMANCE	PnpU
CAIAGVDA and PCAPABILITY	CnpU
CAIAGVDA and MPERFORMANCE	PnmU

CNPL C
Stores the following statistics, depending on the naming convention and type of process stage.

If you specify	Minitab stores
CSIXSIGMA	Cnpl
CAIAGVDA and PPERFORMANCE	PnpL
CAIAGVDA and PCAPABILITY	CnpL
CAIAGVDA and MPERFORMANCE	PnmL

CNPK C

Stores the following statistics, depending on the naming convention and type of process stage.

If you specify	Minitab stores
CSIXSIGMA	Cnpk
CAIAGVDA and PPERFORMANCE	Pnpk
CAIAGVDA and PCAPABILITY	Cnpk
CAIAGVDA and MPERFORMANCE	Pnmk

PPML C (C C)

Stores the observed parts per million or percent less than the lower specification limit. The units match the output. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

PPMU C (C C)

Stores the observed parts per million or percent greater than the upper specification limit. The units match the output. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

PPMT C (C C)

Stores the total parts per million or percent beyond the specification limit. The units match the output. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

ATCAPA: Session command for performing an automated capability analysis

ATCAPA [C E]

Generates a process capability report that automatically selects whether to use a distribution, a transformation, or to make no parametric assumptions about the data. If a data set is arranged in a column, specify the column with subgroup size in E. E can be a column (if the subgroup size varies) or a constant (if the subgroups are the same size.) Subgroup columns may contain numeric, text or date/time data. If your data are in subgroups across rows of several of columns, use RSUBS. For details, go to How to enter subgroup data on page 1246.

The report includes a histogram and the capability statistics for the method that the analysis selected. The report also includes statistics of the process data, such as the mean, median, target (if you enter one), process specifications, and the observed performance. Use the histogram to visually assess whether the process is capable of meeting the specifications consistently.

RSUBS C...C
Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUB, then the subgroup size is 1.

USPEC K
Specifies the upper specification limit. Every capability analysis requires at least 1 specification limit.

LSPEC K
Specifies the lower specification limit. Every capability analysis requires at least 1 specification limit.

Methods

The ATCAPA command has 3 modes.

Specify only the AUTO subcommand
Using the same order as in the list of subcommands and in the table of results, the analysis displays results for the first method that provides reasonable results. If none of the distributions or transformations provides a reasonable fit for the data, then the analysis provides results for a nonparametric method. The analysis also displays a table of statistics about the other methods.

Specify only a subcommand for a method
The analysis displays capability results for the method.

Specify the AUTO subcommand and a subcommand for a method
The analysis displays capability results for the method and a table of statistics about the other methods. Specify only one subcommand for a distribution or transformation.

AUTO
Automatically select a method to use to calculate the capability statistics.

NORMAL
Calculate alternative results with the normal distribution.

WEIBULL
Calculate alternative results with the Weibull distribution.

LNORMAL
Calculate alternative results with the lognormal distribution.

SEXTREMEVALUE
Calculate alternative results with the smallest extreme value distribution.

LEXTREMEVALUE
Calculate alternative results with the largest extreme value distribution.

GAMMA
Calculate alternative results with the gamma distribution.

LOGISTIC
Calculate alternative results with the logistic distribution.

LLOGISTIC
Calculate alternative results with the loglogistic distribution.

EXPONENTIAL
Calculate alternative results with the exponential distribution.

TWEIBULL
Calculate alternative results with the 3-parameter Weibull distribution.

TLNORMAL
Calculate alternative results with the 3-parameter lognormal distribution.

TGAMMA
Calculate alternative results with the 3-parameter gamma distribution.

TLLOGISTIC
Calculate alternative results with the 3-parameter loglogistic distribution.

TEXPONENTIAL
Calculate alternative results with the 2-parameter exponential distribution.

BOXCOX [K]
Calculate alternative results with the Box-Cox transformation. Specify K to provide λ for the transformation. When you do not specify λ, then the analysis calculates the optimal λ value.

JOHNSON [K] [K] [K] [K] [K]
Calculate alternative results with the Johnson transformation. To have the analysis fit the Johnson transformation, do not enter any arguments.

To perform the analysis for a specific Johnson Transformation, specify the arguments. The first argument specifies the family of the transformation. Specify the family with one of the following numbers.

- 1: S_B
- 2: S_L
- 3: S_U

The last 4 arguments provide the coefficients for the transformation function in the following order:

1. γ
2. η
3. ε
4. λ

Insert a value of λ only when the transformation family is S_B or S_U. The following table summarizes the transformation families:

Johnson family	Transformation function	Range
S_B	$\gamma + \eta \ln [(x - \varepsilon) / (\lambda + \varepsilon - x)]$	$\eta, \lambda > 0, -\infty < \gamma < \infty, -\infty < \varepsilon < \infty, \varepsilon < x < \varepsilon + \lambda$
S_L	$\gamma + \eta \ln (x - \varepsilon)$	$\eta > 0, -\infty < \gamma < \infty, -\infty < \varepsilon < \infty, \varepsilon < x$
S_U	$\gamma + \eta \operatorname{Sinh}^{-1} [(x - \varepsilon) / \lambda]$, where $\operatorname{Sinh}^{-1}(x) = \ln [x + \operatorname{sqrt} (1 + x^2)]$	$\eta, \lambda > 0, -\infty < \gamma < \infty, -\infty < \varepsilon < \infty, -\infty < x < \infty$

NPAR

Calculate alternative results with a nonparametric method. Nonparametric methods do not make assumptions about the distribution of the data.

Capability calculation method

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX

Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX

Calculates the capability indices using the ISO method, also called the geometric method.

Options

TARGET K

Specifies a process target value. The results include the target. If the analysis selects a method that uses the target to calculate a capability statistic, like Cpm, then the results include such statistics.

TOLER K

Specifies a value to use in the estimation of the spread of the process. Specify a value such that $0 < K \leq 8$. The application of the value differs slightly among the methods. For example, for an analysis that uses the normal distribution, 6 specifies that the process spread is an interval 6 standard deviations wide, 3 on either side of the process mean. Other methods use intervals that produce comparable measures of capability.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

ALPHA K

Specifies the significance level for the selection of the method. Specify a value such that 0.01 ≤ K ≤ 0.25. By default, the analysis uses a significance level of 0.05. The analysis selects the first distribution or transformation that produces a p-value greater than the significance level. The analysis selects a method even when the method does not produce an exact p-value, such as when the result is that the p-value > 0.25. If no parametric method achieves the significance level, the analysis uses a nonparametric method.

PERCENT

Replaces the default parts per million (PPM) calculations with percentages.

PPM

Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

CONFIDENCE K

Specifies a confidence level in K. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ONECI

Specifies one-sided confidence intervals. The direction of the intervals depends on the statistic. If a larger value indicates a more capable process, then the results include a lower bound. For example, the results include a lower bound for Pp.

If a lower value indicates a more capable process, then the results include an upper bound. For example, the results include an upper bound for the total parts per million out of specification.

TWOCI

Specifies two-sided confidence intervals.

NOCI

Specifies not to include confidence intervals in the output.

Graphs

NOCHART

Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different

file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

The ATCAPA command does not store all of the statistics for a method. To store more statistics, use a capability analysis command that specifies the method.

NAME C
Stores the name of the data columns used in the analysis.

USL C
Stores the upper specification limit.

STARGET C
Stores the process target.

LSL C
Stores the lower specification limit.

MEAN C
Stores the mean of the data.

MEDIAN C
Stores the median of the data.

N C
Stores the number of sample observations.

SDT C
Stores the within-subgroup standard deviation when the method uses the normal distribution or the Box-Cox transformation. Stores a missing for the other methods.

SDO C
Stores the overall standard deviation.

PPMT C
Stores the total parts per million beyond the specification limits or the percent outside of the specification limits. The units depend on the specifications for the analysis.

DISTRIBUTION C
Stores the name of the method that the analysis selects.

CP C
Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cp
CAIAGVDA	Cw

CPK C
Stores the following statistics, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cpk
CAIAGVDA	Cwk

PP C
Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Cnp
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Pnp
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or Cnp
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Pnm
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Pnp
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or Cnp
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Pnm

PPK C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk or Cnpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z or Pnpk
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z or Cnpk
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z or Pnmk
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G or Pnpk
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G or Cnpk
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G or Pnmk

MCAPA: Session command for performing a normal capability analysis for multiple variables

MCAPA C...C

Generates a process capability report when each of the continuous variables follows a normal distribution. You can also use the BOXCOX or JOHNSON subcommand to try to transform your data to follow normal distribution. MCAPA allows you to compare the results for multiple variables or when a grouping variable is associated with a continuous variable. For example, you want to compare the width of parts from the same process before and after a process improvement.

Displays a capability histogram overlaid with a normal curve and displays capability statistics. The distribution curve is generated from the distribution parameters.

The command is entered differently depending on how the data is structured in the worksheet. For details, go to How to enter subgroup data on page 1246. The report that MCAPA produces a histogram, overlaid with a normal curve generated from the process mean and standard deviation. It also displays the locations of the process mean and target, the natural process tolerance, and the process specifications. So you can use the chart to visually assess whether the data are normally distributed, whether the process is centered on the target, and whether it is capable of meeting the process specifications consistently.

If you have data in subgroups, MCAPA uses the pooled standard deviation to estimate the standard deviation by default. Use the RBAR subcommand to base the estimate on subgroup ranges. If you use RBAR, the maximum subgroup size allowed is 100.

For individual observations, the default estimate of sigma is based on the average of the moving range of length 2. Use the MMR subcommand to use the median of the moving range. Use the RSPAN subcommand to change the length of the moving range. The maximum length allowed is 50. Use the SRMSSD subcommand to estimate the standard deviation using the square root of half the mean of the squared successive differences.

BY C...C

Specifies the grouping columns in C...C. Columns may contain numeric or text data. If you use BY subcommand you can only enter only one column with MCAPA command.

BETWEEN

Specifies the use of both between and within subgroup variations for analysis. You cannot use this subcommand with a Johnson transformation.

Subgroup size

Specifies the subgroup size of variables specified with MCAPA command or groups in BY subcommand. You can enter one subgroup size applicable to all variables, one value entered as a constant or in a row of a column for each variable, or one subgroup ID column for each variable.

SIZE K
Specifies the subgroup size for all variables or groups as K

SIZE K...K
Specifies the subgroup size for each variables or groups as K...K

SIZE C
Specifies the subgroup ID column C for all variables or groups

SIZE C...C
Specifies the subgroup ID column for each multiple variables or groups in C...C

Specification limits

Use USPEC and LSPEC to specify a value for the upper and lower specification limits, respectively. You must enter at least one of these.

Use the optional UBOUND and LBOUND to specify the limits as boundaries. K or the values in C are 1 or 0, indicating whether the corresponding limits are boundaries.

If all data sets share the same specification limit, specify a constant or a column with one value. If the data sets have different specification limits, specify a value for each data set in a column or as constants.

USPEC K

USPEC K...K

USPEC C
Specifies the upper specification limit.

UBOUND [K]
Specifies the upper specification limit as the boundary as K, applicable to all variables or groups. K takes values 1 or 0.

UBOUND [K...K]
Specifies the upper specification limit as the boundary as K, one K for each variable or group. K takes values 1 or 0.

UBOUND [C]
Specifies the upper specification limit as the boundary in C. Values in C are 1 and/or 0.

LSPEC K

LSPEC K...K

LSPEC C
Specifies the lower specification limit.

LBOUND [K]
Specifies the lower specification limit as the boundary as K, applicable to all variables or groups. K takes values 1 or 0.

LBOUND [K...K]
Specifies the lower specification limit as the boundary as K, one K for each variable or group. K takes values 1 or 0.

LBOUND [C]
Specifies the lower specification limit as the boundary in C. Values in C are 1 and/or 0.

Historical estimates

It is generally assumed that the data from the process follows a normal distribution. MU is the mean of this distribution. BSIGMA is the process standard deviation between batches and WSIGMA is the process standard deviation within subgroups. You can use these subcommands to enter known values for the process parameters, or enter previous estimates. For BSIGMA, the BETWEEN subcommand must be issued.

You can specify the values in a column or as constants. one value applicable to all variables or groups or one value for each variable or group as a list of constants or a column.

If all data sets share the same process parameter values, specify a constant or a column with one value. If the data sets have different process parameter values, specify a value for each data set in a column or as constants.

You cannot use these subcommands with a transformation.

MU K

MU K...K

MU C
Specifies the mean of the normal distribution.

WSIGMA K

WSIGMA K...K

WSIGMA C
Specifies the historical values for the within-subgroup standard deviation.

BSIGMA K

BSIGMA K...K

BSIGMA C
Specifies the historical values for the between-subgroup standard deviation.

Standard deviation estimates

These subcommands provide options for estimating the within-subgroup process standard deviation based on the following methods.

AMR
Average moving range.

RSPAN K
Use RSPAN to change the length of the moving range. The maximum length allowed is 100.

RBAR
Averages of the subgroup ranges. The maximum subgroup size allowed is 100.

SBAR
Averages of the subgroup standard deviations.

MMR
Median of the moving range.

RSPAN K
Use RSPAN to change the length of the moving range. The maximum length allowed is 100.

SRMSSD
Square root of half of the mean of the squared successive differences.

POOLED
Pooled standard deviation.

BIASED
Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations. The default is to use unbiasing constants.

UNBIASED
Specifies to use unbiasing constants when calculating the within subgroup standard deviations.

OBIASED
Specifies to not use an unbiasing constant when calculating the overall standard deviation.

OUNBIASED
Specifies to use an unbiasing constant when calculating the overall standard deviation.

Box-Cox transformation

BOXCOX [K]

BOXCOX [K...K]

BOXCOX [C]
Transforms the data to follow a normal distribution using Box-Cox transformation. You can also specify a value for lambda to be used in the transformation. If no value is specified, Minitab estimates the optimal value of lambda from the data.

If all data sets share the same lambda value, specify a constant (K) or a column (C) with one value. If the data sets have different lambda values, specify a value for each data set in a column or as constants. K and values in C must lie between -5 and 5.

These subcommands cannot be used with a Johnson transformation.

Johnson transformation

JOHNSON
Selects a distribution function from Johnson system to transform each data set to follow normal distribution. These subcommands cannot be used with a Box-Cox transformation.

PVALUE K...K

PVALUE C

Use the optional PVALUE to change the default p-value of 0.10 to select the optimal transformation. You can specify p-values in a column or as constants, one value for each data set used in the analysis. These subcommands cannot be used with a Box-Cox transformation.

Options

TARGET K

TARGET K...K

TARGET C

Specifies a process target value. Minitab then calculates Cpm, in addition to the standard capability statistics. However, if JOHNSON is issued, then Minitab only displays the target in the process data table and capability histogram.

If all data sets share the same target value, specify a constant or a column with one value. If the data sets have different target values, specify a value for each data set in a column or as constants.

TOLER K

TOLER K...K

TOLER C

By default, calculates the capability statistics using an interval that is 6 standard deviations wide, 3 on either side of the process mean. TOLER lets you change the width of this interval. For example, TOLER 12 says to use an interval 12 standard deviations wide, 6 on either side of the process mean. This is referred to as "6-sigma quality.

If all data sets share the same tolerance, specify a constant or a column with one value. If the data sets have different tolerance, specify a value for each data set in a column or as constants.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX

Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX

Calculates the capability indices using the ISO method, also called the geometric method.

CONFIDENCE [K]

Specifies a confidence level other than the default value of 95.0. Specify any value between 0 and 100. Minitab calculates the confidence intervals for Cp, Cpk, Pp, Ppk, and Cpm statistics.

ONECI

Specifies one-sided confidence intervals. The direction of the intervals depends on the statistic.

Lower bound:

- Cp
- Cpk
- Ppk
- Cpm
- Z.bench

Upper bound:

- Parts Per Million (PPM) < Lower Specification Limit (LSL)
- PPM > Upper Specification Limit (USL)
- PPM total
- Percent < LSL
- Percent > USL
- Percent total

TWOCI

Specifies two-sided confidence intervals.

NOCI

Specifies not to include confidence intervals in the output.

WITHIN

Use WITHIN to perform a within capability analysis only. This subcommand cannot be used with a Johnson transformation.

The default is to conduct both within and overall capability analyses.

OVERALL

Use OVERALL to perform an overall capability analysis only. If a Johnson transformation is used, only an overall analysis can be done so this subcommand is redundant.

The default is to conduct both within and overall capability analyses.

PERCENT
Replaces the default parts per million (PPM) calculations with percentages.

PPM
Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

ZBENCH
Displays benchmark Z statistics rather than the default capability statistics.

CSTAT
Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

Results

BRIEF K
Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	No output is displayed.
2	All output is displayed.

Graphs

GHISTOGRAM
Displays a histogram with a density function.

DSCALING
Use DSCALING to have different X-axis scaling for all the variables.

MINIMUM K

MINIMUM K...K

MINIMUM C
Use MINIMUM and MAXIMUM to enter the range in which the data are plotted in the histogram. If no value is entered, the range for each variable is determined by its data.

MAXIMUM K

MAXIMUM K...K

MAXIMUM C
Use MINIMUM and MAXIMUM to enter the range in which the data are plotted in the histogram. If no value is entered, the range for each variable is determined by its data.

GNORMAL
Displays a normal probability plot for each variable or group.

Storage

NAME C
Stores the names of the data columns used in the analysis or the levels of the by variable that define the different variables.

USL C
Stores the upper specification limit for each variable.

STARGET C
Stores the process target for each variable.

LSL C
Stores the lower specification limit for each variable.

MEAN C
Stores the mean of the data for each variable.

N C
Stores the number of sample observations for each variable.

SDB C
Stores the between-subgroup standard deviation for each variable.

SDW C
Stores the within-subgroup standard deviation for each variable.

SDT C
Stores the within standard deviation for each variable when the analysis is performed for within subgroups variation (default) or stores the total standard deviation for each variable when the analysis is performed for between and within subgroup variation.

SDO C
Stores the overall standard deviation for each variable.

CP C
Stores the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cp or Z.Bench
CAIAGVDA	Cw or Zw.Bench

CPUC C
Stores the upper limit of a confidence interval for the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCp or Z.Bench
CAIAGVDA	UBCw or Zw.Bench

CPLC C

Stores the lower limit of a confidence interval for the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCp or Z.Bench
CAIAGVDA	LBCw or Zw.Bench

CPU C

Stores the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	CPU or Z.USL
CAIAGVDA	CwU or Zw.USL

CPL C

Stores the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	CPL or Z.LSL
CAIAGVDA	CwL or Zw.LSL

CPK C

Stores the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	Cpk
CAIAGVDA	Cwk

CPKU C

Stores the upper limit of a confidence interval for the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	UBCpk
CAIAGVDA	UBCwk

CPKL C

Stores the lower limit of a confidence interval for the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	LBCpk
CAIAGVDA	LBCwk

CCPK C

Stores the following statistics for each variable, depending on the naming convention.

If you specify	Minitab stores
CSIXSIGMA	CCpk
CAIAGVDA	CCwk

PP C

Stores the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Zm.Bench

PPUC C

Stores the upper limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPm.G or Zm.Bench

PPLC C

Stores the lower limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPm.Z or Zm.Bench

If you specify	Minitab stores
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPm.G or Zm.Bench

PPU C

Stores the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPU or Z.USL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpU.Z or Zp.USL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpU.Z or CZp.USL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmU.Z or Zm.USL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpU.G or Zp.USL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpU.G or CZp.USL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmU.G or Zm.USL

PPL C

Stores the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPL or Z.LSL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpL.Z or Zp.LSL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpL.Z or CZp.LSL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmL.Z or Zm.LSL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpL.G or Zp.LSL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpL.G or CZp.LSL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmL.G or Zm.LSL

PPK C

Stores the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G

PPKU C

Stores the upper limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmk.G

PPKL C

Stores the lower limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBPpk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmk.G

CPM C

Stores the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Cpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmm.G

UCPM C

Stores the upper limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	UBCpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	UBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	UBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	UBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	UBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	UBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	UBPmm.G

LCPM C

Stores the lower limit of a confidence interval for the following statistics for each variable, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	LBCpm
CAIAGVDA, PPERFORMANCE, and MTBINDEX	LBPpm.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	LBCpm.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	LBPmm.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	LBPpm.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	LBCpm.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	LBPmm.G

PPML C

Stores the observed parts per million less than the lower specification limit for each variable.

PPMU C

Stores the observed parts per million greater than the upper specification limit for each variable.

PPMT C

Stores the total parts per million beyond the specification limit for each variable.

EPLO C (C C)

Stores either the expected parts per million or the expected percentage less than the lower specification limit for each variable. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUO C (C C)

Stores either the expected parts per million or the expected percentage greater than the upper specification limit for each variable. The units match the output. The calculation uses the overall variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTO C (C C)

Stores either the expected parts per million or the expected percentage outside the specification limits for each variable. The units match the output. The calculation uses the overall variation. Minitab stores the value in the

first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPLW C (C C)
Stores either the expected parts per million or the expected percentage less than the lower specification limit for each variable. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPUW C (C C)
Stores either the expected parts per million or the expected percentage greater than the upper specification limit for each variable. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

EPTW C (C C)
Stores either the expected parts per million or the expected percentage outside the specification limits for each variable. The units match the output. The calculation uses the within subgroup variation. Minitab stores the value in the first column that you specify. To store the confidence bound from the output, specify a second column. To store the confidence interval from the output, specify 2 additional columns.

JTYPE C
Stores the optimal distribution function type selected for Johnson transformation for each variable.

JGAMMA C
Stores the gamma parameter of a distribution function in Johnson transformation for each variable.

JETA C
Stores the eta parameter of a distribution function in Johnson transformation for each variable.

JLAMBDA C
Stores the lambda parameter of a distribution function in Johnson transformation for each variable.

JEPSILON C
Stores the epsilon parameter of a distribution function in Johnson transformation for each variable.

SLAMBDA C
Stores the value of lambda used in Box-Cox transformation for each variable.

TUSL C
Stores the Box-Cox transformed value of the upper specification limit for each variable.

TTARGET C
Stores the Box-Cox transformed target value for each variable.

TLSL C
Stores the Box-Cox transformed value of the lower specification limit for each variable.

TMEA C
Stores the mean of the Box-Cox transformed data for each variable.

TSDB C
Stores the Box-Cox transformed value of between-subgroup standard deviation for each variable.

TSDW C
Stores the Box-Cox transformed value of within-subgroup standard deviation for each variable.

TSDT C
Stores the total standard deviation of the Box-Cox transformed data for each variable.

TSDO C
Stores the overall standard deviation of the Box-Cox transformed data for each variable.

MNCAPA: Session command for performing nonnormal capability analysis for multiple variables

MNCAPA C...C
Generates a process capability report when each of the continuous variables follows a nonnormal distribution or to compare the results for multiple variables when a grouping variable is associated with a continuous variable. For example, you want to compare the width of parts from the same process before and after a process improvement.

The command is entered differently depending on how the data is structured in the worksheet. For details, go to How to enter subgroup data on page 1246. The report includes a capability histogram overlaid with a distribution curve, and a table of overall capability statistics. The distribution curve is generated from the distribution parameters. The report also includes statistics of the process data, such as the mean, parameters, target (if you enter one), and process specifications; the overall capability; and the observed and expected overall performance. So the report can be used to visually assess the distribution of the process relative to the target, whether the data follow a nonnormal distribution, and whether the process is capable of meeting the specifications consistently.

When using the nonnormal model, Minitab calculates the overall capability statistics, Pp, Ppk, PPU, and PPL. The calculations are based on maximum likelihood estimates of the parameters for the specified distribution, rather than mean and variance estimates as in the normal case.

BY C...C
Specifies the grouping columns in C...C. If you use this subcommand, only one variable can be specified with MNCAPA.

USPEC K

USPEC K...K

USPEC C
Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

UBOUND [K]

UBOUND [K...K]

UBOUND [C]
Specifies that the upper specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

LSPEC K

LSPEC K...K

LSPEC C

Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LBOUND [K]

LBOUND [K...K]

LBOUND [C]

Specifies that the lower specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

Distributions

Use the following subcommands to specify which distributions to fit your data with.

EXPONENTIAL

The exponential distribution.

TEXPONENTIAL

The 2-parameter exponential distribution.

LNORMAL

The lognormal distribution.

TLNORMAL

The 3-parameter lognormal distribution.

WEIBULL

The Weibull distribution.

TWEIBULL

The 3-parameter Weibull distribution.

LEXTREMEVALUE

The largest extreme value distribution.

SEXTREMEVALUE

The smallest extreme value distribution.

GAMMA

The gamma distribution.

TGAMMA

The 3-parameter gamma distribution.

LOGISTIC

The logistic distribution.

LLOGISTIC
The loglogistic distribution.

TLLOGISTIC
The 3-parameter loglogistic distribution.

Parameters

SETS K

SETS K...K

SETS C
Sets the shape parameter for the gamma and the Weibull distributions and the scale parameter for all other distributions. Minitab calculates maximum likelihood estimates from the data for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

SETTHRESHOLD K

SETTHRESHOLD K...K

SETTHRESHOLD C
Sets the threshold parameter for the 2-parameter exponential and any 3-parameter distributions. Minitab calculates maximum likelihood estimates from the data set for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

HPARAMETERS K...K

HPARAMETERS C

HPARAMETERS C...C
Sets all the parameters in a distribution. If all data sets share the same parameter values, enter constants or a column. The number of constants or values in a column must equal the number of parameter in a distribution. If the data set has different parameter values, enter columns equal to the number of data sets in your analysis. The number of entries in each column must equal the number of parameters in a distribution.

You must enter parameter values in the following order: location, shape, scale, and threshold. For example, for 3-parameter loglogistic distribution, enter values of the parameters in the following order: location, scale, and threshold.

Capability calculation method

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX
Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX
Calculates the capability indices using the ISO method, also called the geometric method.

Options

TARGET K

TARGET K...K

TARGET C

Specifies a process target value. Minitab then calculates Cpm, in addition to the standard capability statistics.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

TOLER K

TOLER K...K

TOLER C

Specifies a value to calculate the capability statistics using an interval other than 6 standard deviations wide, 3 on either side of the process mean. For example, `TOLER 12` says to use an interval 12 standard deviations wide, 6 on either side of the process mean. This is referred to as "6-sigma quality."

You can specify the values in a column or as constants, one value for each data set used in the analysis.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

ZBENCH

Displays benchmark Z statistics rather than the default capability statistics.

CSTAT

Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

PERCENT

Replaces the default parts per million (PPM) calculations with percentages.

PPM

Replaces the percentage calculations with parts per million (PPM). Use this command to get PPM calculations after you change the default calculations to percentages.

Graphs

GHISTOGRAM

Displays a histogram with a density function for each variable.

DSCALING

Use DSCALING to have different X-axis scaling for all the variables.

MINIMUM K

MINIMUM K...K

MINIMUM C

Use MINIMUM and MAXIMUM to enter the range in which the data are plotted in the histogram. If no value is entered, the range for each variable is determined by its data.

If all the data sets share the same p-value, specify a constant or a column with one value. If the data sets have different p-values, specify a value for each data set in a column or as constants.

MAXIMUM K

MAXIMUM K...K

MAXIMUM C

Use MINIMUM and MAXIMUM to enter the range in which the data are plotted in the histogram. If no value is entered, the range for each variable is determined by its data.

If all the data sets share the same p-value, specify a constant or a column with one value. If the data sets have different p-values, specify a value for each data set in a column or as constants.

GPPLOT

Displays a probability plot for the specified distribution.

NOCHART

Suppresses the display of the graph. The default is to display the graph.

WTITLE

WTITLE specifies the command title of the resulting graph.

Storage

NAME C

Stores the names of the data columns used in the analysis or the levels of the by variable that define the different variables.

USL C

Stores the upper specification limit for each variable.

STARGET C
Stores the process target for each variable.

LSL C
Stores the lower specification limit for each variable.

MEAN C
Stores the mean of the data for each variable.

N C
Stores the number of sample observations for each variable.

SSHAPE C
Stores the shape parameter of a distribution for each variable.

SSCALE C
Stores the scale parameter of a distribution for each variable.

SLOCATION C
Stores the location parameter of a distribution for each variable.

STHRESHOLD C
Stores the threshold parameter of a distribution for each variable.

PP C
Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Pp or Z.Bench
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Pp.Z or Zp.Bench
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cp.Z or CZp.Bench
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pm.Z or Zm.Bench
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Pp.G or Zp.Bench
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cp.G or CZp.Bench
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pm.G or Zm.Bench

PPU C
Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPU or Z.USL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpU.Z or Zp.USL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpU.Z or CZp.USL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmU.Z or Zm.USL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpU.G or Zp.USL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpU.G or CZp.USL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmU.G or Zm.USL

PPL C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	PPL or Z.LSL
CAIAGVDA, PPERFORMANCE, and MTBINDEX	PpL.Z or Zp.LSL
CAIAGVDA, PCAPABILITY, and MTBINDEX	CpL.Z or CZp.LSL
CAIAGVDA, MPERFORMANCE, and MTBINDEX	PmL.Z or Zm.LSL
CAIAGVDA, PPERFORMANCE, and ISOINDEX	PpL.G or Zp.LSL
CAIAGVDA, PCAPABILITY, and ISOINDEX	CpL.G or CZp.LSL
CAIAGVDA, MPERFORMANCE, and ISOINDEX	PmL.G or Zm.LSL

PPK C

Stores the following statistics, depending on the naming convention and calculation method.

If you specify	Minitab stores
CSIXSIGMA	Ppk
CAIAGVDA, PPERFORMANCE, and MTBINDEX	Ppk.Z
CAIAGVDA, PCAPABILITY, and MTBINDEX	Cpk.Z
CAIAGVDA, MPERFORMANCE, and MTBINDEX	Pmk.Z
CAIAGVDA, PPERFORMANCE, and ISOINDEX	Ppk.G
CAIAGVDA, PCAPABILITY, and ISOINDEX	Cpk.G
CAIAGVDA, MPERFORMANCE, and ISOINDEX	Pmk.G

PPML C

Stores the observed parts per million less than the lower specification limit for each variable.

PPMU C

Stores the observed parts per million greater than the upper specification limit for each variable.

PPMT C

Stores the total parts per million beyond the specification limit for each variable.

EPLO C

Stores either the expected parts per million or the expected percentage less than the lower specification limit for each variable. The units match the output. The calculation uses the overall variation.

EPUO C

Stores either the expected parts per million or the expected percentage greater than the upper specification limit for each variable. The units match the output. The calculation uses the overall variation.

EPTO C

Stores either the expected parts per million or the expected percentage outside the specification limits for each variable. The units match the output. The calculation uses the overall variation.

BWSIXPAC: Session command for Between/Within Capability Sixpack

BWSIXPAC C...C *[E]*

Creates a capability report based on the assumption that the data are from normal distribution using both between-subgroup and within-subgroup standard deviations. The report includes the following:

- The individuals, moving range, and R or S charts, to verify whether the process is in control.
- The histogram and normal probability plot, to verify that the data are normally distributed.
- The capability plot, to give you a graphical view of the process variability compared to the specifications. Minitab displays the values to the right of this plot.
- The between/within and overall capability statistics, to help you assess whether a process is in control and is capable of producing output that consistently meets the specifications.

If a data set is arranged in a column, specify the column with BWSIXPAC with subgroup size in E. E can be a column (if the subgroup size varies) or a constant (if the subgroups are the same size.) Columns may contain numeric or text data. If your data are in subgroups across rows of several of columns, use RSUBS. For more information, go to How to enter subgroup data on page 1246.

Minitab determines which chart to display based on the subgroup size. For individual observations, an I-MR chart is displayed. For subgroup size between 2 and 8, an Xbar-R chart is displayed. For subgroup size > 8, an Xbar-S chart is displayed.

RSUBS C...C [K]

Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUBS, then the subgroup size is 1. The optional K indicates the number of data sets if greater than 1. For more information, go to How to enter subgroup data on page 1246.

USPEC C

USPEC K...K

Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

UBOUND C

UBOUND K...K

Specifies that the upper specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

LSPEC C

LSPEC K...K

Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LBOUND C

LBOUND K...K

Specifies that the lower specification is a bound. K or the values in C are 1 or 0, which indicates whether the corresponding limit is a boundary.

Historical estimates

It is generally assumed that the data from the process follow a normal distribution. MU is the mean of this distribution. BSIGMA is the process standard deviation between batches and WSIGMA is the process standard deviation within subgroups. You can use these subcommands to enter known values for the process parameters, or enter previous estimates. If these subcommands are omitted, Minitab estimates the values from the data set.

You can specify the values in a column or as constants, one for each data set used in the analysis.

MU C

MU K...K
Specifies historical values for the mean of the normal distribution.

BSIGMA C

BSIGMA K...K
Specifies historical values for the between-subgroup standard deviation.

WSIGMA C

WSIGMA K...K
Specifies historical values for the within-subgroup standard deviation.

Options

TARGET C

TARGET K...K
Specifies the target values for each data set.

TOLER K
Specifies to calculate the capability statistics using an interval specified as K.

CSIXSIGMA
Specifies the Six Sigma capability naming conventions. `CSIXSIGMA` matches the naming convention used in Minitab version 22.5 and before.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA
Specifies the AIAG/VDA capability naming conventions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE
Specifies the capability indices for process performance.

PCAPABILITY
Specifies the capability indices for long-term capability.

MPERFORMANCE
Specifies the capability indices for machine performance.

The following subcommands specify the method to calculate the capability indices.

MTBINDEX
Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX
Calculates the capability indices using the ISO method, also called the geometric method.

CSTAT
Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

ZBENCH
Displays benchmark Z statistics rather than the default capability statistics.

Box-Cox transformation

BOXCOX [C]

BOXCOX K...K
Transforms the data to follow a normal distribution using Box-Cox transformation. You can also specify a value for lambda to be used in the transformation. If no value is specified, Minitab estimates the optimal value of lambda from the data.

You can specify the values for lambda in a column (C) or as constants (K), one value for each data set used in the analysis. K and values in C must lie between −5 and 5.

Standard deviation estimates for capability statistics

OBIASED
Specifies to not use an unbiasing constant when calculating the overall standard deviation.

OUNBIASED
Specifies to use an unbiasing constant when calculating the overall standard deviation.

AMR
Estimates the between-subgroup standard deviation using the average moving range method using the length specified with RSPAN as K.

MMR
Estimates the between-subgroup standard deviation using the median of the moving ranges method using the length specified with RSPAN as K.

SRMSSD
Estimates the between-subgroup standard deviation using the square root of half of the mean of the squared successive differences.

RBAR
Estimates the within-subgroup standard deviation using the average of subgroup range method.

SBAR
Estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

POOLED
Estimates the within-subgroup standard deviation using the pooled standard deviation.

Standard deviation estimates for control charts

CCAMR
Estimates standard deviation between subgroups using the average moving range method using the length specified with RSPAN as K.

CCMMR
Estimates standard deviation between subgroups using the median of the moving ranges method using the length specified with RSPAN K.

CCRBAR
When the subgroup size <= 8, CCRBAR estimates the within-subgroup standard deviation using the average of subgroup range method.

CCSSBAR
When the subgroup size <= 8, CCSSBAR estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCSPOOLED
When the subgroup size <= 8, CCSPOOLED estimates the within-subgroup standard deviation using the pooled standard deviation.

CCSBAR
When the subgroup size > 8, CCSBAR estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCPOOLED
When the subgroup size > 8, CCPOOLED estimates the within-subgroup standard deviation using the pooled standard deviation.

Standard deviation estimates for both capability statistics and control charts

RSPAN K
Specifies the length of the moving range for AMR, MMR, CCAMR, and CCMMR.

BIASED
Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations.

UNBIASED
Specifies to use unbiasing constants when calculating the within subgroup standard deviations.

Tests for special causes

TEST K...K

Specifies one or more of the eight tests for special causes, where K takes values 1, 2, 3, 4, 5, 6, 7, and 8. Each K corresponds to a test.

Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

When any point fails one of the eight tests, Minitab displays the information in the output.

K	Test
1	One point more than 3 standard deviation from center line
2	Nine points in a row on same side of center line
3	Six points in a row, all increasing or all decreasing
4	Fourteen points in a row, alternating up and down
5	Two out of three points more than 2 standard deviation from center line (same side)
6	Four out of five points more than 1 standard deviation from center line (same side)
7	Fifteen points in a row within 1 standard deviation of center line (either side)
8	Eight points in a row more than 1 standard deviation from center line (either side)

Graphs

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

SIXPACK: Session command for **Normal Capability Sixpack**

SIXPACK [C...C *E*]
Generates a capability report based on the assumption that the data are from normal distribution. You can also use the BOXCOX or JOHNSON subcommand to try to transform your data to follow a normal distribution. The report includes the following:

- X, R, and run charts, to verify that the process is in a state of control.
- A histogram and a normal probability plot, to verify that the data are normally distributed.
- A capability plot, for a graphical view of the process variability compared to the specifications.
- Capability statistics Cp, CPU, CPL, Cpk, and benchmark Z values, to help you assess whether a process is in control and is capable of producing output that consistently meets the specifications.

If a data set is arranged in a column, specify the column with SIXPACK with subgroup size in E. E can be a column (if the subgroup size varies) or a constant (if the subgroups are the same size.) If your data are in subgroups across rows of several of columns, use RSUBS. For more information, go to How to enter subgroup data on page 1246.

By default, Minitab determines which control charts to display based on the subgroup size. For individual observations, an I chart and an MR chart are displayed. For subgroup size between 2 and 8, an Xbar chart and an R chart are displayed. For subgroup size greater than 8, an Xbar chart and an S chart are displayed. The CCRBAR and CCSBAR subcommands can be used to force the display of an R chart or an S chart, respectively, for all subgroup sizes greater than 1.

RSUBS C...C [K]
Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUB, then the subgroup size is 1. The optional K indicates the number of data sets if greater than 1. For more information, go to How to enter subgroup data on page 1246.

USPEC C

USPEC K...K

Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LSPEC C

LSPEC K...K

Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

TARGET C

TARGET K...K

Specifies a process target value. Minitab then calculates Cpm, in addition to the standard capability statistics. You can specify values in a column or as constants, one value for each data set used in the analysis.

Historical estimates

It is generally assumed that the data from the process follow a normal distribution. MU is the mean of this distribution. BSIGMA is the process standard deviation between batches and WSIGMA is the process standard deviation within subgroups. You can use these subcommands to enter known values for the process parameters, or enter previous estimates. If these subcommands are omitted, Minitab estimates the values from the data set.

You can specify the values in a column or as constants, one for each data set used in the analysis.

You cannot use these subcommands with JOHNSON or BOXCOX.

MU C

MU K...K

Specifies historical values for the mean of the normal distribution.

SIGMA C

SIGMA K...K

Specifies historical values for the standard deviation of the normal distribution.

Options

TEST K...K

Specifies one or more of the eight tests for special causes, where K takes values 1, 2, 3, 4, 5, 6, 7, and 8. Each K corresponds to a test.

Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated.

When any point fails one of the eight tests, Minitab displays the information in the output.

K	Test
1	One point more than 3 standard deviation from center line

K	Test
2	Nine points in a row on same side of center line
3	Six points in a row, all increasing or all decreasing
4	Fourteen points in a row, alternating up and down
5	Two out of three points more than 2 standard deviation from center line (same side)
6	Four out of five points more than 1 standard deviation from center line (same side)
7	Fifteen points in a row within 1 standard deviation of center line (either side)
8	Eight points in a row more than 1 standard deviation from center line (either side)

BREAKOUT K

Specifies the number of subgroups or observations to display in the run chart. The default number is 25.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. `CSIXSIGMA` matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability index naming convention.

You cannot use `CAIAGVDA` with `CSIXSIGMA`. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX

Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX

Calculates the capability indices using the ISO method, also called the geometric method.

CSTAT

Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

ZBENCH

Displays benchmark Z statistics rather than the default capability statistics.

Box-Cox transformation

BOXCOX [C]

BOXCOX [K...K]

Transforms the data to follow a normal distribution using Box-Cox transformation. You can also specify a value for lambda to be used in the transformation. If no value is specified, Minitab estimates the optimal value of lambda from the data.

You can specify the values for lambda in a column (C) or as constants (K), one value for each data set used in the analysis. K and values in C must lie between −5 and 5.

You cannot use BOXCOX with JOHNSON.

Johnson transformation

JOHNSON

Selects a distribution function from the Johnson system to transform data in each column to follow a normal distribution.

You cannot use JOHNSON with BOXCOX.

PVALUE K...K

PVALUE C

Use the optional PVALUE to change the default p-value of 0.10 to select the optimal transformation. You can specify p-values in a column or as constants, one value for each data set used in the analysis.

Standard deviation estimates for capability statistics

OBIASED

Specifies to not use an unbiasing constant when calculating the overall standard deviation.

You can use this subcommand with JOHNSON.

OUNBIASED

Specifies to use an unbiasing constant when calculating the overall standard deviation.

You can use this subcommand with JOHNSON.

AMR

When the subgroup size = 1, AMR estimates the between-subgroup standard deviation using the average moving range method using the length specified with RSPAN as K.

You cannot use this subcommand with JOHNSON.

MMR

When the subgroup size = 1, MMR estimates the between-subgroup standard deviation using the median of the moving ranges method using the length specified with RSPAN as K.

You cannot use this subcommand with JOHNSON.

SRMSSD

When the subgroup size = 1, SRMSSD estimates the between-subgroup standard deviation using the square root of half of the mean of the squared successive differences.

You cannot use this subcommand with JOHNSON.

RBAR

When the subgroup size > 1, RBAR estimates the within-subgroup standard deviation using the average of subgroup range method.

You cannot use this subcommand with JOHNSON.

SBAR

When the subgroup size > 1, SBAR estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

You cannot use this subcommand with JOHNSON.

POOLED

When the subgroup size > 1, POOLED estimates the within-subgroup standard deviation using the pooled standard deviation.

You cannot use this subcommand with JOHNSON.

Standard deviation estimates for control charts

CCAMR

Estimates the between-subgroup standard deviation using the average moving range method using the length specified with RSPAN as K.

CCMMR

Estimates the between-subgroup standard deviation using the median of the moving ranges method using the length specified with RSPAN K.

CCRBAR

When the subgroup size <= 8, CCRBAR estimates the within-subgroup standard deviation using the average of subgroup range method.

CCSSBAR

When the subgroup size <= 8, CCSSBAR estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCSPOOLED

When the subgroup size <= 8, CCSPOOLED estimates the within-subgroup standard deviation using the pooled standard deviation.

CCSBAR

When the subgroup size > 8, CCSBAR estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCPOOLED

When the subgroup size > 8, CCSPOOLED estimates the within-subgroup standard deviation using the pooled standard deviation.

Standard deviation estimates for both capability statistics and control charts

RSPAN K
Specifies the length of the moving range for AMR, MMR, CCAMR, and CCMMR.

BIASED
Specifies to not use unbiasing constants when calculating the within-subgroup standard deviations.

UNBIASED
Specifies to use unbiasing constants when calculating the within subgroup standard deviations.

Graphs

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

NNSIXPACK: Session command for **Nonnormal Capability Sixpack**

NNSIXPACK C...C
Generates a capability report based on the assumption that the data follows a nonnormal distribution. The report combines the following information into a single display:

- An Xbar chart (or I chart for individual observations)
- An R or S chart (or MR chart for individual observations)
- A run chart of the last 25 subgroups (or last 25 observations)
- A histogram of the process data
- A distribution probability plot
- A process capability plot
- Overall capability statistics Pp, Ppk, and distribution parameters

The Xbar, R, and run charts can be used to verify that the process is in a state of control. The histogram and distribution probability plot can be used to verify that the data approximate a specified distribution. Lastly, the capability plot gives a graphical view of the process variability compared to the specifications. Combined with the capability statistics, this information can help you assess whether your process is in control and can produce output that consistently meets the specifications.

When using the nonnormal model, Minitab only calculates the overall capability statistics, Pp and Ppk. The calculations are based on maximum likelihood estimates of the distribution parameters for the specified distribution, rather than mean and variance estimates as in the normal case.

RSUBS C...C [K]
Specifies that the data in each row of the input columns makes up a subgroup. All columns must be the same length. If only one input column is given with RSUB, then the subgroup size is 1.

The optional K indicates the number of data sets if greater than 1. For details, go to How to enter subgroup data on page 1246.

USPEC C

USPEC K...K
Specifies the upper specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

LSPEC C

LSPEC K...K

Specifies the lower specification limit.

Important You must enter either an upper specification limit in USPEC or a lower specification limit in LSPEC.

Distributions

Use the following subcommands to specify which distributions to fit your data with.

EXPONENTIAL

The exponential distribution.

TEXPONENTIAL

The 2-parameter exponential distribution.

LNORMAL

The lognormal distribution.

TLNORMAL

The 3-parameter lognormal distribution.

FNORMAL

The folded normal distribution.

WEIBULL

The Weibull distribution.

TWEIBULL

The 3-parameter Weibull distribution.

LEXTREMEVALUE

The largest extreme value distribution.

SEXTREMEVALUE

The smallest extreme value distribution.

GAMMA

The gamma distribution.

TGAMMA

The 3-parameter gamma distribution.

LOGISTIC

The logistic distribution.

LLOGISTIC

The loglogistic distribution.

TLLOGISTIC

The 3-parameter loglogistic distribution.

Parameters

SETS K...K

SETS C

Sets the shape parameter for the gamma and the Weibull distributions and the scale parameter for all other distributions. Minitab calculates maximum likelihood estimates from the data for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

SETTHRESHOLD K...K

SETTHRESHOLD C

Sets the threshold parameter for the 2-parameter exponential and any 3-parameter distributions. Minitab calculates maximum likelihood estimates from the data set for missing values.

You can specify values for the parameters in a column or as constants, one value for each data set used in the analysis.

HPARAMETERS K...K

HPARAMETERS C

HPARAMETERS C...C

Sets all the parameters in a distribution. If all data sets share the same parameter values, enter constants or a column. The number of constants or values in a column must equal the number of parameter in a distribution. If the data set has different parameter values, enter columns equal to the number of data sets in your analysis. The number of entries in each column must equal the number of parameters in a distribution.

You must enter parameter values in the following order: location, shape, scale, and threshold. For example, for 3-parameter loglogistic distribution, enter values of the parameters in the following order: location, scale, and threshold.

Standard deviation estimates for control charts

RSPAN K

Specifies the length of the moving range for CCAMR and CCMMR.

BIASED

Specifies to not use unbiasing constants when calculating the within subgroup standard deviations.

UNBIASED

Specifies to use unbiasing constants when calculating the within subgroup standard deviations.

Standard deviation estimates for control charts, when subgroup size = 1

CCAMR

Estimates the within-subgroup standard deviation using the average moving range method using the length specified with RSPAN as K.

CCMMR
Estimates the within-subgroup standard deviation using the median of the moving ranges method using the length specified with RSPAN K.

Standard deviation estimates for control charts, when 1 < subgroup size <= 8

CCRBAR
Estimates the within-subgroup standard deviation using the average of subgroup range method.

CCSSBAR
Estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCSPOOLED
Estimates the within-subgroup standard deviation using the pooled standard deviation.

Standard deviation estimates for control charts, when subgroup size > 8

CCSBAR
Estimates the within-subgroup standard deviation using the average of the subgroup standard deviation method.

CCPOOLED
Estimates the within-subgroup standard deviation using the pooled standard deviation.

Capability calculation method

The following subcommands specify the method to calculate the capability indices. You can specify only one subcommand. If you do not specify a subcommand, Minitab uses the default setting from preferences.

MTBINDEX
Calculates the capability indices using the MTB method, also called the Z-score method.

ISOINDEX
Calculates the capability indices using the ISO method, also called the geometric method.

Options

BREAKOUT K
Specifies the number of subgroups or observations to display in the run chart. The default number is 25.

TOLER K...K

TOLER C
Specifies a value to calculate the capability statistics using an interval other than 6 standard deviations wide, 3 on either side of the process mean. For example, `TOLER 12` says to use an interval 12 standard deviations wide, 6 on either side of the process mean. This is referred to as "6-sigma quality."

You can specify the values in a column or as constants, one value for each data set used in the analysis.

CSIXSIGMA

Specifies the Six Sigma capability index naming convention. CSIXSIGMA matches the naming convention used in Minitab 22.5 and earlier versions.

You cannot use CAIAGVDA with CSIXSIGMA. If you do not specify either subcommand, Minitab uses the default setting from preferences.

CAIAGVDA

Specifies the AIAG/VDA capability naming convention.

You cannot use CAIAGVDA with CSIXSIGMA. If you do not specify either subcommand, Minitab uses the default setting from preferences.

The following subcommands specify the type of capability indices. If you do not specify a subcommand, Minitab uses the default setting from preferences.

PPERFORMANCE

Specifies the capability indices for process performance.

PCAPABILITY

Specifies the capability indices for long-term capability.

MPERFORMANCE

Specifies the capability indices for machine performance.

CSTAT

Displays the capability statistics when the default preference is set to display the benchmark Z statistics.

ZBENCH

Displays benchmark Z statistics rather than the default capability statistics.

Graph

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

BCAPA: Session command for performing binomial capability analysis

BCAPA C...C *E*
Generates a process capability report for data from a binomial distribution. E is the sample size for each data set, which can be a column (if subgroup size varies) or a constant (if all samples are the same size). For more information, go to How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART on page 1244.

The report includes the following:

- A P chart, to verify that the process was in a state of control at the time the report was generated.
- A chart of cumulative %defective, to verify that you collected data from enough samples to have a stable estimate of the %defective.
- A histogram of %defective, to display the overall distribution of the %defectives from the samples collected.
- A defective rate plot, to verify that the %defective is not influenced by the number of items that were sampled.

If an observation is missing, there is a gap in the P chart where that subgroup would have been plotted. The other plots and charts simply exclude the missing observation.

In the P chart, the control limits are a function of the subgroup size. In general, the control limits are further from the centerline for smaller subgroups than they are for larger ones. When you have unequal subgroup sizes, the plot of %defective versus sample size helps you verify that there is no relationship between the two. For example,

if you tend to have a smaller %defective when more items are sampled, this could be caused by fatigued inspectors, a common problem. The subgroup size has no bearing on the other charts, since they display only the %defective.

TARGET C

TARGET K...K

Specifies a process target value, which can be in a column or as list of constants. The number of rows in a column and constants must equal the number of columns listed with the main command.

P C

P K...K

Specify historical values for the proportions of defectives, which can be in a column or as list of constants. The number of rows in a column and constants must equal the number of columns listed with the main command.

TEST K...K

Specifies one or more of the four tests for special causes, where K takes values 1, 2, 3, and 4. Each value of K corresponds to a test.

Value of K	Test
1	One point more than 3 standard deviations from center line
2	Nine points in a row on same side of center line
3	Six points in a row, all increasing or all decreasing
4	Fourteen points in a row, alternating up and down

Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated. When any point fails one of the four tests, Minitab displays the information in the output.

Graphs

NOCHART

Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*

You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different

file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores column names used in the analysis.

PAVERAGE C
Stores the average proportion of defectives.

PALC C
Stores the lower limit of the confidence interval for the average proportion of defectives.

PAUC C
Stores the upper limit of the confidence interval for the average proportion of defectives.

PDEFECTIVE C
Stores the percent defective.

PDLC C
Stores the lower limit of the confidence interval for the percent defectives.

PDUC C
Stores the upper limit of the confidence interval for the percent defectives.

PDTARGET C
Stores the target percent of defectives.

PPM C
Stores the parts per million defective.

PPML C
Stores the lower limit of the confidence interval for the parts per million defectives.

PPMU C
Stores the upper limit of the confidence interval for the parts per million defectives.

ZPROCESS C
Stores the process Z.

ZPLC C
Stores the lower limit of the confidence interval for the process Z.

ZPUC C
Stores the upper limit of the confidence interval for the process Z.

PCAPA: Session command for performing a Poisson capability analysis

PCAPA C...C *E*
Generates a process capability report for data from a Poisson distribution. E is the subgroup size for each data set, which can be a column (if subgroup size varies) or a constant (if the subgroups are the same size). For more information, go to How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART on page 1246.

The report includes:

- A U chart, to verify that the process was in a state of control at the time the report was generated. Note: if the samples are all the same size, there are two cases to consider. First, if all the samples have more than 1 unit of size. In that case, a U chart is still preferred because it charts the number of defects per unit, whereas a C chart plots the number of defects only. Defects per unit is more informative than just defects. If the samples all have 1 unit of size, then a U chart and a C chart are exactly the same.
- A chart of cumulative mean DPU, to verify that you collected data from enough samples to have a stable estimate of the mean.
- A histogram of DPU, to display the overall distribution of the defects per unit from the samples collected.
- A defect rate plot, to verify that DPU is not influenced by the size of the items sampled.

If an observation is missing, there is a gap in the U chart where that subgroup would have been plotted. The other plots and charts simply exclude the missing observation.

In the U chart, the control limits are a function of the subgroup size. In general, the control limits are further from the centerline for smaller subgroups than they are for larger ones. When you do have unequal subgroup size, the plot of DPU versus sample size will permit you to verify that there is no relationship between the two. For example, if you tend to have a smaller DPU when more items are sampled, this could be caused by fatigued inspectors, a common problem. The subgroup size has no bearing on the other charts, since they only display the DPU.

TARGET C

TARGET K...K
Specifies the target values for each data set.

Historical estimates

MU C

MU K...K
The process is assumed to produce data from a population that follows a Poisson distribution. The mean and variance of this distribution are both m. MU specifies a value for this parameter. You can specify the values for this parameter in a column or as list of constants, where each values corresponds to the columns listed in the main command.

Tests for special causes

TEST K
Specifies one or more of the four tests for special causes, where K takes values 1, 2, 3, and 4. Each value of K corresponds to a test.

Value of K	Test
1	One point more than 3 standard deviations from center line
2	Nine points in a row on same side of center line
3	Six points in a row, all increasing or all decreasing
4	Fourteen points in a row, alternating up and down

Each test detects a specific pattern in the data plotted on the chart. The occurrence of a pattern suggests a special cause for the variation, one that should be investigated. When any point fails one of the four tests, Minitab displays the information in the output.

Graph

NOCHART
Suppresses the display of the graph. The default is to display the graph.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

WTITLE *"title"*
You can use WTITLE as a subcommand with LAYOUT and all graphs. The title that you specify becomes the command title of the resulting graph.

GSAVE *"file_name"*

GSAVE K
Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

Storage

NAME C
Stores the names of the data columns used in the analysis.

DEFECTS C
Stores the mean number of defects.

DEFL C
Stores the lower limit of the confidence interval for the mean number of defects.

DEFU C
Stores the upper limit of the confidence interval for the mean number of defects.

DPU C
Stores the mean defects per unit.

DPUL C
Stores the lower limit of the confidence interval for the mean defects per unit.

DPUU C
Stores the upper limit of the confidence interval for the mean defects per unit.

DMINIMUM C
Stores the minimum number of defects per unit.

DMAXIMUM C
Stores the maximum number of defects per unit.

DTARGET C
Stores the target defects per unit.

Tolerance Intervals

TOLINTERVALS: Session command for calculating tolerance intervals

TOLINTERVALS C...C

TOLINTERVALS K K K
Computes a range of values for a specific quality characteristic of a product that likely covers a specified proportion of future product output. Alternatively, a lower or upper limit may be constructed such that the specified proportion will be greater than or less than the limit. Tolerance intervals specify both a population proportion and a confidence level. They are frequently used to detect excessive variation by comparing client requirements to tolerance limits that cover a proportion of the population that is acceptable to the manufacturer.

TOLINTERVALS C...C calculates exact parametric and nonparametric tolerance intervals for each variable C...C. TOLINTERVALS K K K calculates tolerance intervals using summarized values for the sample size, mean, and standard deviation.

Options

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PPERCENT K
Specifies the minimum percentage of the population covered by the tolerance interval. K is typically any number in the interval [0.000001, 99.9999]. However, if K < 0.999999, then a population percentage of K*100 is assumed. The default value is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Graphs

GTIPLOT

Displays a tolerance interval plot for each sample. Each plot consists of a histogram and two interval plots based on the normal and nonparametric tolerance intervals.

Results

Each of the following subcommands displays a table in the output. If you do not use any of these subcommands, all the output is displayed.

TMETHOD

Displays the method table that contains the hypotheses and significance level.

TSTATISTICS

TSTATISTICS displays the sample size and point estimates of the mean and standard deviation for each population.

TTOLINTERVAL

TTOLINTERVAL displays both the normal and nonparametric tolerance intervals.

Storage

SPTLIMITS C C

SPTLIMITS stores the limits for the tolerance intervals that are calculated using the normal method.

The lower limits is stored in the first column C and the upper limit is stored in the second column C. If ITYPE is specified as –1 or 1, then specify only one column C.

SNPTLIMITS C C

SNPTLIMITS stores the limits for the tolerance intervals that are calculated using the nonparametric method.

The lower limits is stored in the first column C and the upper limit is stored in the second column C. If ITYPE is specified as –1 or 1, then specify only one column C.

STFACTOR C

Stores the tolerance factor for the normal method, also known as the k factor, for each variable. The values are stored in column C.

NNTINTERVALS: Session command for calculating tolerance intervals

NNTINTERVALS C...C

Computes a range of values for a specific quality characteristic of a product that likely covers a specified minimum percentage of current or future product output. Alternatively, a lower or upper limit may be constructed such that the specified proportion will be greater than or less than the limit. Tolerance intervals specify both a population proportion and a confidence level. They are frequently used to detect excessive variation by comparing client requirements to tolerance limits that cover a proportion of the population that is acceptable to the manufacturer.

NNTINTERVALS C...C calculates exact parametric and nonparametric tolerance intervals for each variable C...C.

PPERCENT K

Specifies the minimum percentage of the population covered by the tolerance interval. K is typically any number in the interval [0.000001, 99.9999]. However, if K < 0.999999, then a population percentage of K*100 is assumed. The default value is 95.

Distributions

EXPONENTIAL

Assumes that the data follow an exponential distribution.

LNORMAL

Assumes that the data follow a lognormal distribution.

WEIBULL

Assumes that the data follow a Weibull distribution.

SEXTREMEVALUE

Assumes that the data follow a smallest extreme value distribution.

LEXTREMEVALUE

Assumes that the data follow a largest extreme value distribution.

GAMMA

Assumes that the data follow a gamma distribution.

LOGISTIC

Assumes that the data follow a logistic distribution.

LLOGISTIC

Assumes that the data follow a loglogistic distribution.

NOPARA

Assumes the data do not follow any specific distribution. Minitab provides only a nonparametric interval.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Graphs

GTIPLOT

Displays a tolerance interval plot for each sample. Each plot consists of a histogram and two interval plots based on the parametric and nonparametric tolerance intervals.

Results

Each of the following subcommands displays a table in the output. If you do not use any of these subcommands, all the output is displayed.

TMETHOD

Displays the method table that contains the hypotheses and significance level.

TSTATISTICS

TSTATISTICS displays the sample size and point estimates of the mean and standard deviation for each population.

TESTIMATE

TESTIMATE displays the estimated values of the distribution parameters.

TTOLINTERVAL

TTOLINTERVAL displays both the parametric and nonparametric tolerance intervals.

Storage

SPTLIMITS C C

SPTLIMITS stores the limits for the tolerance intervals that are calculated using the parametric method.

The lower limits is stored in the first column C and the upper limit is stored in the second column C. If ITYPE is specified as –1 or 1, then specify only one column C.

SNPTLIMITS C C

SNPTLIMITS stores the limits for the tolerance intervals that are calculated using the nonparametric method.

The lower limits is stored in the first column C and the upper limit is stored in the second column C. If ITYPE is specified as –1 or 1, then specify only one column C.

STFACTOR C
Stores the tolerance factor for the parametric method, also known as the k factor, for each variable. The values are stored in column C.

Measurement Systems Analysis

TOGAGE: Session command for performing a Type 1 gage study

TOGAGE C
Performs a Type 1 gage study on measurements from a single part. The argument specifies the numeric column.

Reference value

REFERENCE K
Specifies the reference value for the single part. Reference value must be numeric.

Tolerance limits

Important You must use one of the following tolerance limit subcommands.

TOLERANCE K (K)
Specifies the tolerance range, which is USL–LSL or USL–0, when both limits are known or a natural zero exists. You must specify tolerance to calculate the capability indices.

USL K
Use when only the upper specification limit is known. Minitab does not calculate capability indices unless both specifications are given.

LSL K
Use when only the lower specification limit is known. Minitab does not calculate capability indices unless both specifications are given.

Gage information

GAGE *"text"*
Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K
Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*
Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*

Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

CGCONSTANT K

Specifies the percent of tolerance used for calculating the capability indices, Cg and CgK. K is a constant and must be a positive number. The default value of K is 20.

STUDYVAR K

Specifies the constant to multiply by sigma for the gage variation. K is a constant and must be a positive number. The default value of K is 6.

RESOLUTION K

Specifies the gage resolution value. K is a constant and must be a positive number. If you do not use this subcommand, Minitab does not display the resolution output.

NOBIASTEST

Suppresses the testing of bias=0.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

VDA 5

VDA5

Displays results for VDA 5, such as uMS and %QMS. Specify subsubcommands to include additional sources of uncertainty in the calculations. Specify uncertainty values greater than 0.

CALIBRATION K

Specifies the standard deviation of measurements that differ because of the calibration for the reference standard (uCAL). Usually, this value comes from the calibration certificate.

LINEARITY K

Specifies the standard deviation of measurements that differ because of linearity (uLIN). Usually, this value comes from a linearity study where the reference part was within the range of measurements.

REST K...K

REST C

Specifies the uncertainty due to other factors (uREST). Specify up to 100 values that are greater than 0. You can also specify a column from the worksheet. In the column, use up to 100 values that are greater than 0.

GAWORKSHEET: Session command for creating a gage R&R study worksheet

GAWORKSHEET K K K

Creates a Gage R&R Study Worksheet (or Attribute Agreement Analysis Worksheet) with Parts (Samples), Operators (Appraisers), Run Order, and Standard Order columns. The first K is the number of parts (samples). The second K is the number of operators (appraisers). The third K is the number of replicates.

PARTS K...K

PARTS *"text"..."text"*

PARTS C

Specifies part names for all K parts, the first number specified with GAWORKSHEET. By default, the part names are "1", "2", ... "K". If you use a column of part names, each value must be distinct.

OPERATORS K...K

OPERATORS *"text"*

Specifies operator names for all K operators, the second number specified with GAWORKSHEET. By default, the operator names are "1", "2", ... "K".

Options

NSTANDARD K...K

NSTANDARD C

Specifies the numeric standard values for samples in Attribute Agreement Analysis. Arguments must be numeric values and the number of standard values must be equal to the number of samples specified with GAWORKSHEET. When the standards are stored in a column, the number of values must equal the number of samples. This subcommand cannot be used if TSTANDARD has been issued.

REFERENCE K

Specifies a reference level out of the two standard values to define the value order for the stored binary standard value column. This subsubcommand is only valid with binary data. If the standard has more than two levels, REFERENCE is ignored.

TSTANDARD *"text"..."text"*

TSTANDARD C

Specifies the text standard values for samples in Attribute Agreement Analysis. Arguments must be text values and the number of standard values must be equal to the number of samples specified with GAWORKSHEET. When the standards are stored in a column, the number of values must equal the number of samples. This subcommand cannot be used if NSTANDARD has been issued.

REFERENCE *"text"*

Specifies a reference level out of the two standard values to define the value order for the stored binary standard value column. This subsubcommand is only valid with binary data. If the standard has more than two levels, REFERENCE is ignored.

RANDOMIZE (K)

Completely randomizes the run order. To generate the same random sample, you can optionally set the base, K. K is a numeric value.

WRANDOMIZE (K)

Randomizes the run order within each operator. To generate the same random sample, you can optionally set the base, K. K is a numeric value.

RSOPERATOR

When randomizing the run order within operator (WRANDOMIZE), this subcommand randomizes the order of operators. Without this subcommand, Operator 1 is always listed first, Operator 2 is second, and so on.

SORDER C

Stores the standard order in the worksheet. You can designate the column to store it in and name the column.

RORDER C

Stores the run order in the worksheet. You can designate the column to store it in and name the column.

XMATRIX C C (C)

Stores the part and operator names in the worksheet. Optionally, stores standard values for each sample in an Attribute Agreement Analysis worksheet. You can designate and name the storage columns.

Results

RTABLE

Displays a summary table in the output.

RNGAGE: Session command for creating gage run chart

RNGAGE

Plots the response values by operator and part number.

You can use this chart to quickly assess the differences in measurements between different operators and parts. A stable process would give you a random horizontal scattering of points. With an operator or part effect, you would see some kind of pattern or patterns in the plot.

A horizontal reference line is drawn at the mean of all observations. Use the MU subcommand to enter your own value for the reference line.

Type the command as shown with example data in C1–C3.

```
RNGAGE;
  PARTS C1;
  OPERS C2;
  RESPONSE C3.
```

Measurement data

PARTS C (required)

Specifies the part names or numbers. Part numbers can be text or numeric.

OPERS C

Specifies the operator names or numbers. Operators can be text or numeric.

TRIALS C

Specifies trial numbers.

A gage R&R study usually consists of several operators taking a series of measurements on each of several parts with a common gage. It is often recommended that the order of runs in the gage R&R experiment be randomized to reduce the effects of systematic errors. If TRIAL is not used, the first value encountered in each part by operator combination is taken to be the first trial, the second part encountered with the same part and operator is taken to be the second trial, and so on. If the actual order of the trials is different, you should use TRIAL to designate the order. When designating the order, specify the order within the particular set of part and operator values, not the order within the whole experiment.

RESPONSE C (required)
Specifies the observed measurements.

MU K
Specifies a value for the horizontal reference line on the Gage Run Chart. By default, the reference line is the mean of all observations.

Gage information

GAGE *"text"*
Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K
Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*
Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

GTOL *"text"*
Specifies the gage tolerance in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*
Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MGAGE: Session command for performing an expanded gage R&R study

MGAGE C = Terms
Use an expanded gage R&R study when you want to assess the variation due to Part, Operator (optional), and a maximum of eight additional factors (optional). All factors can be fixed or random, crossed or nested, and text or numeric. The design can be balanced or unbalanced.

This method breaks down the gage variation into reproducibility and repeatability. You may also add other main and interaction effects to Reproducibility in addition to, or instead of, Operator and Operator by Part. Optionally, you can specify terms in addition to Part to calculate part-to-part variation.

The main command to perform a gage R&R study for multiple factors is MGAGE C = Terms, where C is the numeric measurement column (response). Both the response and term columns must have the same length. For more information, go to How to specify the model in MGAGE on page 1253.

Measurement data

PARTS C (required)

Specifies the part names or numbers. Part numbers can be text or numeric. The column must be specified by MGAGE.

OPERATORS C (optional)

Specifies the operator names or numbers. Operators can be text or numeric. If used, the column must be specified by MGAGE.

Gage information

GAGE *"text"*

Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K

Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*

Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*

Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

FIXED C...C (optional)

Specifies which factors are fixed. By default, all factors are random. If used, the column or columns must be specified by MGAGE. Do not include interaction terms or nested factors.

VARPART *terms* (optional)

Specifies the terms to define Part-to-Part variation. Without this subcommand, the Part-to-Part variation is defined by Part term alone. Operator term is not allowed. If used, the column or columns must be specified by MGAGE.

STUDYVAR K

Specifies the multiple used in the Study Var (K * SD) column of the output. The default multiple is 6, which is the number of standard deviations needed to capture 99% of your process measurements.

SIGMA K

Specifies a historical standard deviation (σ).

VHIST

Estimate the total process variation with the historical standard deviation. When you specify a historical standard deviation, Minitab uses this option by default.

VSTUDY

Calculate the total process variation with the parts in the study. The output table that displays the percentage contributions will include a %Process column when you use this subcommand. This column displays the percentage of the process (σ) that is taken up by each variance component.

TOLER K

Specifies a tolerance range. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

LSL K

Specifies the lower specification limit when no upper specification limit exists. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

USL K

Specifies the upper specification limit when no lower specification limit exists. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

PVALUE K

Specifies the alpha level to determine which interaction terms are insignificant. K is typically any number between 0 and 1. However, if K > 0.999999, alpha = K /100. The default is 0.05.

NOPCONTRIBUTION

Suppresses the %Contribution output and the corresponding bars in the Variance Components graph.

NOPSTUDYVAR

Suppresses the %Study Var output and the corresponding bars in the Variance Components graph.

VDA 5

VDA5

Displays results for VDA 5, such as uMP and %QMP. Specify subsubcommands to include additional sources of uncertainty in the calculations. Specify uncertainty values greater than 0.

RESOLUTION K

Specifies the gage resolution value (uRE). Resolution is the least increment of measurement that the gage accurately reports. For example, if the gage accurately reports measurements to the tenths place, then K = 0.1.

CALIBRATION K

Specifies the uncertainty of measurements that differ because of the calibration for the reference standard (uCAL). Usually, this value comes from the calibration certificate.

LINEARITY K
Specifies the uncertainty of measurements that differ because of linearity (uLIN). Usually, this value comes from a linearity study where the reference part was within the range of measurements.

REST K...K

REST C
Specifies the uncertainty due to other factors (uREST). Specify up to 100 values that are greater than 0. You can also specify a column from the worksheet. In the column, use up to 100 values that are greater than 0.

REPEATABILITY K
Specifies the uncertainty of measurements that differ because the measurement process produces different measurements for the same part (uEVO). Usually, this value comes from the standard deviation of the measurements of a reference part in a Type I Gage Study.

BIAS K
Specifies is the uncertainty in the measurements due to the difference from the known reference measurement and the average of the measurements in the study (uBI). Usually, this value comes from the estimated bias of the measurements in a Type I Gage Study.

Results

TFACTINFO
Displays the factor information table.

TANOVA
Displays ANOVA tables.

TVARCOMP
Displays the variance component table.

TSTUDYVAR
Displays the study variance and tolerance table.

TDCATEGORY
Displays the number of distinct categories.

TRISK
Displays the conditional and joint probability of misclassification tables.

Graphs

GVARCOMP
Displays a graphical representation of the Gage R&R.

GXBARR C [C]
Displays Xbar and R Charts. One or two Cs are required. These columns cannot include Part and all columns must be specified by MGAGE.

GFMAIN C...C
Displays the measurements by single factor plots. At least one C is required and all columns must be specified by MGAGE.

GFINT C...C

Displays two-way interaction plots. Minitab draws a single interaction plot if you enter two factors, or multiple interaction plots if you enter more than two factors. At least two Cs are required and no nested factors are allowed as arguments. All columns must be specified by MGAGE.

SEPARATE

Displays all graphs, except Xbar and R chart, separately. By default, MGAGE produces layouts of 6 graphs in one frame.

Storage

SVARCOMP C C

Stores the variance component and percent contribution for each source listed in an ANOVA table in C and C, respectively.

SSTDDEV C

Stores the standard deviation for each source listed in an ANOVA table in C.

SSTUDYVAR C C

Stores the study variation and the percent study variation for each source listed in an ANOVA table in C and C, respectively.

SPROCESS C

Stores the %Process values in C. This subcommand is ignored if SIGMA is not issued.

STOLERANCE C

Stores the %Tolerance values in C. This subcommand is ignored if TOLER, LSL, or USL is not issued.

LNGAGE: Session command for performing a gage linearity and bias study

LNGAGE

Performs a gage linearity and bias study.

A gage linearity and bias study tells you how accurate your measurements are through the expected range of the measurements. It answers the question, "Does my gage have the same accuracy for all sizes of objects being measured?" A gage accuracy study examines the difference between the observed average measurement and a reference value. A gage accuracy study answers the question, "How accurate is my gage when compared to a master value?"

Both studies are done by selecting parts whose measurements cover the normal range of values for a particular process, measuring the parts with a master system, then having an operator make several measurements on each part using a common gage. Minitab subtracts each measurement taken by the operator from the reference measurement, then calculates, for each part, an average deviation from the master measurement.

To calculate the linearity of the gage, Minitab finds the best-fit line relating the average deviations to the reference measurements. Then,

Linearity = | slope | * process standard deviation

Generally, the closer the slope is to zero, the better the gage linearity. Linearity can also be expressed as a percentage of the process variation by multiplying the slope of the line by 100.

To calculate the accuracy of the gage, Minitab combines the deviations from the reference measurement for all parts. The mean of this combined sample is the gage accuracy. Accuracy can also be expressed as a percentage of the overall process variation by dividing the mean deviation by the process standard deviation, and multiplying by 100.

Type the command as shown with example data in C1-C3.

```
LNGAGE;
  PARTS C1;
  REFERENCE C2;
  RESPONSE C3
  SIGMA 14.1941.
```

Linearity data

PARTS C (required)

Specifies the part names or numbers. Part numbers can be text or numeric.

REFERENCE C (required)

Specifies the reference values.

RESPONSE C (required)

Specifies the observed measurements.

SIGMA K

Specifies a historical standard deviation (σ). The table in the output that displays the percentage contributions will include a %Process column when you use this subcommand. This column displays the percentage of the process (σ) taken up by each variance component.

Gage information

GAGE *"text"*

Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K

Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*

Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

GTOL *"text"*

Specifies the gage tolerance in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*

Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

STDEV
Specifies the sample standard deviation to estimate the repeatability standard deviation. If you do not use STDEV, Minitab estimates the repeatability standard deviation using the sample range.

GAGERR: Session command for performing a crossed gage R&R study

GAGERR
Performs a gage R&R analysis using the ANOVA method or the Xbar and R method.

The Xbar and R method breaks down the gage variation into reproducibility and repeatability. The ANOVA method examines the Operator by Part effect component of Reproducibility.

Measurement data

PARTS C (required)
Specifies the part names or numbers. Part numbers can be text or numeric.

OPERS C
Specifies the operator names or numbers. Operator names can be text or numeric. This subcommand is optional for GAGERR when the ANOVA method is used, but required for GAGERR when XBAR is used.

RESPONSE C (required)
Specifies the observed measurements.

Gage information

GAGE *"text"*
Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K
Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*
Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

GTOL *"text"*
Specifies the gage tolerance in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*

Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

XBAR

Specifies the Xbar and R method, which breaks down the overall variation into three categories: part-to-part, repeatability, and reproducibility. If you do not enter the XBAR subcommand, Minitab uses the ANOVA method. The ANOVA method goes one step further and breaks down reproducibility into its operator and operator-by-part components.

STUDYVAR K

Specifies the multiple used in the Study Var (K * SD) column of the output. The default multiple is 6, which is the number of standard deviations needed to capture 99% of your process measurements.

SIGMA K

Specifies a historical standard deviation (σ).

VHIST

Estimate the total process variation with the historical standard deviation. When you specify a historical standard deviation, Minitab uses this option by default.

VSTUDY

Calculate the total process variation with the parts in the study. The output table that displays the percentage contributions will include a %Process column when you use this subcommand. This column displays the percentage of the process (σ) that is taken up by each variance component.

SPROCESS C

Stores the process variation for each source in an ANOVA table in C.

TOLER C

Specifies a tolerance range. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

STOLERANCE C

Stores the process tolerance for each source in an ANOVA table in C.

LSL K

Specifies the lower specification limit of the process. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

USL K

Specifies the upper specification limit of the process. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

PVALUE K

Specifies the p-value used to eliminate the interaction term from a two-way ANOVA model when the ANOVA method is used to calculate Gage variance components. The default p-value is 0.05.

RISK

Calculates the joint and conditional probabilities of misclassification when an upper specification, lower specification, or both are specified.

SEPARATE

Displays each graph separately. By default, GAGERR produces layouts of 6 graphs in one frame.

TITLE *"text"*

Specifies a title for the graph. When you omit this subcommand, Minitab displays a default title.

NOPCONTRIBUTION

Suppresses the %Contribution output and the corresponding bars in the Variance Components graph.

NOPSTUDYVAR

Suppresses the %Study Var output and the corresponding bars in the Variance Components graph.

VDA 5

VDA5

Displays results for VDA 5, such as uMP and %QMP. Specify subsubcommands to include additional sources of uncertainty in the calculations. Specify uncertainty values greater than 0.

RESOLUTION K

Specifies the gage resolution value (uRE). Resolution is the least increment of measurement that the gage accurately reports. For example, if the gage accurately reports measurements to the tenths place, then K = 0.1.

CALIBRATION K

Specifies the uncertainty of measurements that differ because of the calibration for the reference standard (uCAL). Usually, this value comes from the calibration certificate.

LINEARITY K

Specifies the uncertainty of measurements that differ because of linearity (uLIN). Usually, this value comes from a linearity study where the reference part was within the range of measurements.

REST K...K

REST C

Specifies the uncertainty due to other factors (uREST). Specify up to 100 values that are greater than 0. You can also specify a column from the worksheet. In the column, use up to 100 values that are greater than 0.

REPEATABILITY K

Specifies the uncertainty of measurements that differ because the measurement process produces different measurements for the same part (uEVO). Usually, this value comes from the standard deviation of the measurements of a reference part in a Type I Gage Study.

BIAS K

Specifies is the uncertainty in the measurements due to the difference from the known reference measurement and the average of the measurements in the study (uBI). Usually, this value comes from the estimated bias of the measurements in a Type I Gage Study.

Storage

SVARCOMP C C
Stores the variance component and percent contribution for each source listed in an ANOVA table in C and C.

SSTDEV C
Stores the standard deviation for each source listed in an ANOVA table in C.

SSTUDYVAR C C
Stores the study variation and the percent study variation for each source listed in an ANOVA table in C and C.

SDISTINCT C
Stores the number of distinct categories that the process is able to distinguish in C.

Confidence intervals

CONFIDENCE K
Specifies a confidence level in K. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95. CONFIDENCE is ignored if XBAR is issued.

ITYPE K
Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SIVARCOMP C [C]
Stores the confidence limits for the variance components for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPVARCOMP C [C]
Stores the confidence limits for the %contribution of the variance components for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SISTDDEV C [C]
Stores the confidence limits for the standard deviations for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SISV C [C]
Stores the confidence limits for the study variation for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPSV C [C]
Stores the confidence limits for the %study variation for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SITOLERANCE C [C]
Stores the confidence limits for the %tolerance for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPROCESS C [C]

Stores the confidence limits for the %process values for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIDSTINCT C [C]

Stores the confidence limits for the number of distinct categories in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

WHEELER: Session command for performing a crossed Evaluate Measurement Process (EMP) study

WHEELER

Performs an Evaluate Measurement Process (EMP) study using the ANOVA method or the Xbar and R method. The Xbar and R method breaks down the gage variation into reproducibility and repeatability. The ANOVA method also examines the Operator by Part component of Reproducibility.

Measurement data

PARTS C

Specifies the part names or numbers. Part numbers can be text or numeric.

OPERATORS C

Specifies the operator names or numbers. Operator names can be text or numeric. This subcommand is optional when the ANOVA method is used, but required when XBAR is used.

RESPONSE C

Specifies the observed measurements.

Options

ANOVA

Specifies the ANOVA method, which breaks down the overall variation into four categories: part-to-part, repeatability, reproducibility, and the repeatability by reproducibility. If you do not enter the XBAR subcommand or the ANOVA subcommand, then Minitab uses the ANOVA method.

XBAR

Specifies the Xbar and R method, which breaks down the overall variation into three categories: part-to-part, repeatability, and reproducibility. If you do not enter the XBAR subcommand or the ANOVA subcommand, then Minitab uses the ANOVA method. The ANOVA method goes one step further and breaks down reproducibility into its operator and operator-by-part components.

LSL K

Specifies the lower specification limit of the process. When you specify at least 1 specification limit and use the ANOVA method, then the TRISK subcommand adds a table that shows the probabilities of misclassification.

USL K

Specifies the upper specification limit of the process. When you specify at least 1 specification limit and use the ANOVA method, then the TRISK subcommand adds a table that shows the probabilities of misclassification.

SIGMA K

Specifies a historical standard deviation (σ). The analysis uses the historical standard deviation to calculate probabilities of misclassification. Probabilities of misclassification require at least 1 specification limit and the use of the ANOVA method.

ALPHA K

Specifies the significance level for the decision limits on the ANOME and ANOMR charts. The default value is 0.05. Specify a value between 0 and 0.5.

INCREMENT K

By default, the analysis calculates the measurement increment from the data. Use this subcommand to specify a measurement increment. For example, the gage that you use reports results in the tenths place, but records measurements only to the nearest half, such as 1.0 and 1.5. Then, specify that the increment is 0.5. Specify a positive value.

Graphs

GREPEAT

Displays the repeatability plot of the part measurements. The measurements are by operator if the analysis specifies the operators. When the number of measurements on a single part is 8 or less, the chart displays the ranges of the measurements. For larger sample sizes, the chart displays the standard deviations.

GSIGNAL

Displays the signal-to-noise plot of the mean part measurements. The measurements are by operator if the analysis specifies the operators.

GPARALLELISM

Displays the parallelism plot of the part measurements. The measurements are by operator. The analysis ignores this subcommand if the analysis does not specify the operators.

GVARIANCE

Displays the plot of the contributions of the variance components.

GMAIN

Displays the plot of the main effect of different operators. The analysis ignores this subcommand if the analysis does not specify the operators.

GMRANGE

Displays the plot of the average ranges of different operators. The analysis ignores this subcommand if the analysis does not specify the operators.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TVARCOMP

Displays the variance components table.

TESTAT

Displays the EMP statistics table and the gage classification table.

TRISK
Displays tables that show the conditional and joint probability of misclassification tables. The analysis calculates the probabilities of misclassification when the analysis uses the ANOVA method and the analysis includes at least 1 specification limit.

TRESO
Displays the effective resolution table.

STORAGE

SVARCOMP C
Stores the variance components in C. This statistic is the Variance column from the Variance Components table.

SPCONTRIBUTION C
Stores the percent of the total variation for the variance components in C. This statistic is the %Total column from the Variance Components table.

SSTDDEV C
Stores the standard deviations for the variance components in C. This statistic is the StdDev column from the Variance Components table.

NTGAGE: Session command for performing a nested gage R&R study

NTGAGE
Use a nested gage R&R study to assess the variation in your measurement system when every operator cannot measure all parts.

If you need to use destructive testing, then you must be able to assume that all parts within a single batch are identical enough to claim that they are the same part. If you are unable to make that assumption, then part-to-part variation within a batch will mask the measurement system variation. If you can make that assumption, then your choice of using a crossed gage R&R study or a nested gage R&R study depends on how your measurement process is set up. If all operators measure parts from each batch, then use a crossed or expanded study. If each batch is measured by only a single operator, then use a nested or expanded study. In fact, whenever operators measure unique parts, the design is nested.

Measurement data

PARTS C (required)
Specifies the part names or numbers. Part numbers can be text or numeric.

OPERS C (required)
Specifies the operator names or numbers. Operator names can be text or numeric.

RESPONSE C (required)
Specifies the observed measurements.

Gage information

GAGE *"text"*
Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K
Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*
Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

GTOL *"text"*
Specifies the gage tolerance in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*
Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

STUDYVAR K
Specifies the multiple used in the Study Var (K * SD) column of the output. The default multiple is 6, which is the number of standard deviations needed to capture 99% of your process measurements.

SIGMA K
Specifies a historical standard deviation (σ).

VHIST
Estimate the total process variation with the historical standard deviation. When you specify a historical standard deviation, Minitab uses this option by default.

VSTUDY
Calculate the total process variation with the parts in the study. The output table that displays the percentage contributions will include a %Process column when you use this subcommand. This column displays the percentage of the process (σ) that is taken up by each variance component.

SPROCESS C
Stores the process variation for each source in an ANOVA table in C.

TOLER C
Specifies a tolerance range. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

STOLERANCE C
Stores the process tolerance for each source in an ANOVA table in C.

LSL K

Specifies the lower specification limit when no upper specification limit exists. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

USL K

Specifies the upper specification limit when no lower specification limit exists. The output table that displays the percentage contributions will include a %Tolerance column when you use this subcommand. This column displays the percentage of tolerance taken up by each variance component.

RISK

Calculates the joint and conditional probabilities of misclassification when an upper specification, lower specification, or both are specified.

SEPARATE

Displays each graph separately. By default, NTGAGE produces layouts of 6 graphs in one frame.

TITLE *"text"*

Specifies a title for the graph. When you omit this subcommand, Minitab displays a default title.

NOPCONTRIBUTION

Suppresses the %Contribution output and the corresponding bars in the Variance Components graph.

NOPSTUDYVAR

Suppresses the %Study Var output and the corresponding bars in the Variance Components graph.

VDA 5

VDA5

Displays results for VDA 5, such as uMP and %QMP. Specify subsubcommands to include additional sources of uncertainty in the calculations. Specify uncertainty values greater than 0.

RESOLUTION K

Specifies the gage resolution value (uRE). Resolution is the least increment of measurement that the gage accurately reports. For example, if the gage accurately reports measurements to the tenths place, then K = 0.1.

CALIBRATION K

Specifies the uncertainty of measurements that differ because of the calibration for the reference standard (uCAL). Usually, this value comes from the calibration certificate.

LINEARITY K

Specifies the uncertainty of measurements that differ because of linearity (uLIN). Usually, this value comes from a linearity study where the reference part was within the range of measurements.

REST K...K

REST C

Specifies the uncertainty due to other factors (uREST). Specify up to 100 values that are greater than 0. You can also specify a column from the worksheet. In the column, use up to 100 values that are greater than 0.

REPEATABILITY K

Specifies the uncertainty of measurements that differ because the measurement process produces different measurements for the same part (uEVO). Usually, this value comes from the standard deviation of the measurements of a reference part in a Type I Gage Study.

BIAS K
Specifies is the uncertainty in the measurements due to the difference from the known reference measurement and the average of the measurements in the study (uBI). Usually, this value comes from the estimated bias of the measurements in a Type I Gage Study.

Storage

SVARCOMP C C
Stores the variance components and %contribution in C and C.

SSTDEV C
Stores the standard deviations in C.

SSTUDYVAR C C
Stores the study variation and percent %study variation in C and C.

SDISTINCT C
Stores the number of distinct categories in C.

Confidence intervals

CONFIDENCE K
Specifies a confidence level in K. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95. Displays confidence intervals for the variability measures with a confidence level of K. CONFIDENCE is ignored if XBAR is issued.

ITYPE K
Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

SIVARCOMP C [C]
Stores the confidence limits for the variance components for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPVARCOMP C [C]
Stores the confidence limits for the %contribution of the variance components for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SISTDDEV C [C]
Stores the confidence limits for the standard deviations for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SISV C [C]
Stores the confidence limits for the study variation for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPSV C [C]
Stores the confidence limits for the %study variation for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SITOLERANCE C [C]
Stores the confidence limits for the %tolerance for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIPROCESS C [C]
Stores the confidence limits for the %process values for each source listed in an ANOVA table in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

SIDSTINCT C [C]
Stores the confidence limits for the number of distinct categories in C and C. Use two arguments for a two-sided confidence interval and one for a one-sided interval.

LONGMETHOD: Session command for performing an attribute gage study (analytic method)

LONGMETHOD
Performs an attribute gage study (analytic method).

Attribute gage study data

PARTS C (required)
Specifies the part names or numbers. Part numbers can be text or numeric.

REFERENCE C
Specifies the reference value column for each part. Reference values must be numeric. Structure raw data so that each row contains the part name or number, reference value, and binary response. Structure summarized data so that each row contains the part name or number, reference value, and summarized count.

RESPONSE C (K)

RESPONSE (C)
RESPONSE C specifies the binary column of raw data. The values in this column can be numeric or text.

RESPONSE C (K) or RESPONSE C (C) specifies the summarized count column and the number of times each part is examined (a constant or column). If all parts are examined the same number of times, K represents the number of trials. If the parts are examined a different number of times, you must create a column that contains the number of trials.

ATTRIBUTE K
Specifies the attribute (or standard). This value must be the same data type as the response and must match one of the two distinct response values.

For summarized data, ATTRIBUTE K is optional. If ATTRIBUTE is not specified, the default is "Acceptance".

Tolerance limits

USELOW
Specifies the use of the lower limit to calculate bias and repeatability. You must use either USELOW or USEHIGH.

USEHIGH
Specifies the use of the upper limit to calculate bias and repeatability. You must use either USELOW or USEHIGH.

TOLERANCE K (K)
Specifies the tolerance limit or limits for accepting or rejecting the product.

If one argument is given, K is the tolerance limit corresponding to the USELOW or USEHIGH subcommand. If two arguments are given, the first K is the lower limit and second K is the upper limit. The first K must be less than the second K.

Gage information

GAGE *"text"*
Specifies the name of your gage in the label on the output. If you do not use this subcommand, the label will contain a blank line.

DATE K
Specifies the date that the analysis is performed in the label on the output. K is a constant and a text value. If you do not use this subcommand, the label will contain a blank line.

USER *"text"*
Specifies the user name in the label on the output. If you do not use this subcommand, the label will contain a blank line.

GTOL *"text"*
Specifies the gage tolerance in the label on the output. If you do not use this subcommand, the label will contain a blank line.

MISC *"text"*
Specifies any miscellaneous notes in the label on the output. If you do not use this subcommand, the label will contain a blank line.

Options

REGTEST
Specifies the regression method to test for bias in the attribute gage system. If you do not include this subcommand, the AIAG method is used to test bias. You must have at least 20 trials per part to perform the AIAG method.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

ATGAGE: Session command for performing an attribute agreement analysis

ATGAGE

Use attribute agreement analysis to assess the agreement of nominal or ordinal ratings given by multiple appraisers for both nominal and ordinal data. The measurements are subjective ratings by people rather than direct physical measurements. Examples include:

- Automobile performance ratings
- Classification of fabric quality as good or bad
- Ratings of coffee color, aroma, and taste on a 1–10 scale

In these situations, quality characteristics are difficult to define and assess. To obtain meaningful classifications, more than one appraiser should classify the response measure. If the appraisers agree, the possibility exists that the ratings are accurate. If the appraisers disagree, rating usefulness is limited.

Agreement data

SAMPLES C

Specifies the sample column. Values can be numeric or text. Must be used in conjunction with APPRAISERS subcommand, and data must be stacked.

APPRAISERS C

Specifies the appraiser column. Values can be numeric or text. Must be used in conjunction with SAMPLES subcommand, and data must be stacked.

UNSTACK K K

Specifies that the data are in unstacked format. The first K is the number of appraisers, the second K is the number of trials for each appraiser. Each K must be a positive integer. Cannot be used in conjunction with SAMPLES or APPRAISERS subcommands.

NAMES C

NAMES K K

Specifies the names of the appraisers for unstacked data (the default is Appraiser 1, ..., Appraiser k). The number of rows in the column (or the number of constants) must equal the number of appraisers. The argument can be text or numeric. If no names are specified and there is only one trial per appraiser, the column names are used as the appraiser's names.

This subcommand is ignored if the data are stacked.

RESPONSE C

RESPONSE C...C

Specifies the response column. When used in conjunction with SAMPLES and APPRAISERS (stacked data) a single column must be specified. When used in conjunction with UNSTACK (unstacked), the number of arguments must equal the number of appraisers times the number of trials.

The response column or columns can either be numeric or text.

The design must be balanced. (There must be the same number of trials because each appraiser must evaluate each sample the same number of times.)

For more information, go to Data requirements and formats for attribute agreement analysis on page 1221.

ATTRIBUTE C

Specifies the attribute (or standard) for each sample if known.

Must be the same data type as the response. Missing attributes are not allowed.

ORDINAL

Specifies that the data are ordinal with more than two levels.

To specify the value order for a text response column, right-click the column, then choose **Column Properties** > **Value Order**. Value ordering for all text response columns must be the same.

The value ordering for the attribute (standard) column must be the same as for the response column or columns.

Information about the attribute agreement analysis

DATE *"text"*

Specifies the date that the attribute gage analysis is performed. K is a constant and text value.

USER *"text"*

Specifies the name of the user who does the analysis. K is a constant and text value.

PRODUCT *"text"*

Specifies the product name.

MISC *"text"*

Specifies information that is not included by DATE, USER, and PRODUCT. K is a constant and text value.

Options

COHEN

Calculates Cohen's alpha when appropriate. Minitab will calculate Cohen's kappa when two appraisers rate a single trial or when each appraiser rates two trials.

DISAGREEMENT

Displays the disagreement table. Minitab displays how often each appraiser's assessments differ from each known standard or attribute value. Must be used in conjunction with ATTRIBUTE subcommand.

CONFIDENCE K

Specifies a confidence level in K for the interval estimation of the percentage agreement within appraisers, and between each appraiser and the known standard. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

Graphs

GAPPRAISER

Specifies the graph of the matched proportions for each appraiser when the number of replicates for each appraiser is more than one. If there is no replicate, and GAPPRAISER is entered, no graph is created, and a message of "No graph is made when no repeated ratings are given by each appraiser" will appear in the output.

GATTRIBUTE

Specifies the graph of the matched proportions between the ratings of each appraiser and the attribute when attribute is known. If the attribute is unknown, and GATTRIBUTE is entered, then no graph is created, and a message of "No graph of the matched proportions between the ratings for each appraiser and the attribute is made when the attribute for each sample is unknown." will appear in the output.

ONEDOC

Displays both plots together.

Results

BRIEF K

Controls the amount of output with a value in K.

Value of K	Output that is displayed
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays percentages.
2	Minitab displays percentages, and kappa and Kendall's (ordinal data) coefficient tables.

Acceptance Sampling

AASA: Session command for acceptance sampling by attributes

AASA K

Creates or compares attributes acceptance sampling plans using fraction defective or defects per unit to represent lot/process quality.

Value of K	Representation of lot/process quality
1	Fraction defective
2	Defects per unit

Create plans

Creates a sampling plan based on the criteria you enter.

CREATE

Creates a sampling plan based on plan criteria.

ALPHA K

Specifies the probability that a lot will meet the AQL and will not be accepted, also known as the producer's risk. You can specify K as 0<K<1. The default value of K is 0.05.

UBALPHA K

Specifies the maximum alpha allowed.

BETA K

Specifies the probability that a lot will exceed the RQL and will still be accepted, also known as the consumer's risk. You may specify K as 0<K<1. The default value of K is 0.10.

UBBETA K

Specifies the maximum beta allowed.

Compare plans

Compares several user defined sampling plans.

You can vary sample sizes (SIZE) and/or acceptance numbers (ANUMBER):

- Specify one sample size and several acceptance numbers to examine the effect of varying only the acceptance number.
- Specify multiple sample sizes with only one acceptance number to examine the effect of varying only sample size.
- Specify combinations of sample size and acceptance number to examine specific paired combinations. Use the same number of sample sizes and acceptance numbers.

COMPARE

Compares several user defined sampling plans.

SIZE K...K

Specifies the sample sizes of the plans you wish to compare.

ANUMBER K...K

Specifies the accept number (number of defective units or total defects allowed) of the plans you wish to compare.

Plan criteria

AQL K

Specifies the acceptable quality level. AQL is required when you create plans, but optional when you compare plans.

RQL K

Specifies the rejectable quality level. RQL is required when you create plans, but optional when you compare plans.

LOTSIZE K (optional)

Specifies the lot size. The lot size is required for the ATI and AOQ curves.

HYPERGEOMETRIC

Specifies to use the hypergeometric distribution when creating/comparing sampling plans of isolated lots when you have go / no go defectives data. You must also specify the lot size in this case.

By default, Minitab uses the binomial distribution to create sampling plans and compare sampling plans for go/no go data.

To correctly use the binomial distribution, Minitab assumes that the sample comes from a large lot or from a stream of lots randomly that are selected from an ongoing process. However, if the lot of products that you are sampling from is an isolated lot of finite size—for example, you receive one special order shipment of 500 labels—then the exact distribution for calculating the probability of acceptance is the hypergeometric distribution.

Measurement type

PROPORTION

Choose to use proportion defective or defects per unit to represent the quality of a lot. If you choose proportion, enter the plan AQL and RQL in proportions. For example, AQL = 0.02 and RQL = 0.08.

PERCENT

Choose to use percent defectives or defects per hundred units to represent the quality of a lot. If you choose percent, enter the plan AQL and RQL in percents. For example, AQL = 2.0 and RQL = 8.0.

PPM

Choose to use proportion defectives per million or defects per million units to represent the quality of a lot. If you choose ppm, enter the plan AQL and RQL in ppm. For example, AQL = 200 and RQL = 800.

Graphs

GOC

Creates an operating characteristic (OC) curve, which describes the discriminatory power of an acceptance sampling plan. The OC curve plots the probabilities of accepting a lot versus the fraction defective.

GAOQ

Creates an average outgoing quality (AOQ) curve, which approximates the relationship between the quality of the incoming material and the quality of the outgoing material, assuming that rejected lots will be 100% inspected and defective items will be reworked and inspected again (rectifying inspection).

The average outgoing quality depends on the incoming quality, the probability that the lot will be accepted, and the sample size and the lot size.

GATI

Creates an average total inspection (ATI) curve, which approximates the relationship between the quality of the incoming material and the number of items that need to be inspected, assuming that rejected lots will be 100% inspected and defective items will be reworked and inspected again (rectifying inspection).

The average total inspection per lot depends on the incoming quality, the probability that the lot will be accepted, and the sample size and the lot size.

ONEGRAPH

Displays all graphs on one graph.

Options

DEFECTIVES K...K

Calculates the probability of acceptance for additional defective levels, specified by K...K.

Storage

SSIZE C

Stores the sample size.

SANUMBER C

Stores the accept number.

SDEFECTIVES C

Stores the defectives/defect levels. When you use this command, Minitab stores the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SPACCEPT C

Stores the probability of accepting values. When you use this command, Minitab stores the acceptance values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SPREJECT C

Stores the probability of rejecting values. When you use this command, Minitab stores the acceptance values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES. The probability of rejecting is 1 – the probability of accepting.

SAOQ C

Stores the average outgoing quality values. When you use this command, Minitab stores the AOQ values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SATI C

Stores the average total inspection values. When you use this command, Minitab stores the ATI values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

VASAMPLING: Session command for creating or comparing variables acceptance plans

VASAMPLING

Creates a variables acceptance sampling plan based on the criteria of AQL, RQL, and lot size. This sampling plan tells you how many units to sample and what the accept/reject criteria is in terms of critical distances.

Create plans

Creates a sampling plan based on the criteria you enter.

CREATE

Creates a sampling plan based on plan criteria.

ALPHA K

Specifies the probability that a lot will meet the AQL and will not be accepted. You can specify K as 0<K<1. If you do not specify K, the default is K=0.05.

BETA K

Specifies the probability that a lot will exceed the RQL and still be accepted. You can specify K as 0<K<1. If you do not specify K, the default is K=0.10.

Compare plans

Compares several user defined sampling plans.

You can vary sample sizes (SIZE) and/or critical distances (CDISTANCE):

- Specify one sample size and several critical distances to consider the effect of varying only critical distance.
- Specify multiple sample sizes with only one acceptance number to consider the effect of varying only sample size.

- Specify combinations of sample size and critical distance to consider specific paired combinations. Use the same number of sample sizes and critical distances.

COMPARE
Compares several user defined sampling plans.

SIZE K...K
Specifies the sample sizes of the plans you wish to compare.

CDISTANCE K...K
Specifies the critical distances (criteria to accept or reject a lot) of the plans you wish to compare.

Plan criteria

LSPEC K
Specifies the lower specification limit for the measurements. You must enter at least one specification limit.

USPEC K
Specifies the upper specification limit for the measurements. You must enter at least one specification limit.

SIGMA K
Optionally, specifies the standard deviation.

AQL K
Specifies the acceptable quality level. Is required when creating plans, but optional when comparing plans.

RQL K
Specifies the rejectable quality level. Is required when creating plans, but optional when comparing plans.

LOTSIZE K
Optionally, specifies the lot size. The lot size is required for the ATI and AOQ curves.

Measurement type

PROPORTION
Choose to use proportion defective or defects per unit to represent the quality of a lot. If you choose proportion, enter the plan AQL and RQL in proportions. For example, AQL = 0.02 and RQL = 0.08.

PERCENT
Choose to use percent defectives or defects per hundred units to represent the quality of a lot. If you choose percent, enter the plan AQL and RQL in percents. For example, AQL = 2.0 and RQL = 8.0.

PPM
Choose to use proportion defectives per million or defects per million units to represent the quality of a lot. If you choose ppm, enter the plan AQL and RQL in ppm. For example, AQL = 200 and RQL = 800.

Graphs

GOC
Creates an operating characteristic (OC) curve, which describes the discriminatory power of an acceptance sampling plan. The OC curve plots the probabilities of accepting a lot versus the fraction defective.

GAOQ

Creates an average outgoing quality (AOQ) curve, which approximates the relationship between the quality of the incoming material and the quality of the outgoing material, assuming that rejected lots will be 100% inspected and defective items will be reworked and inspected again (rectifying inspection).

The average outgoing quality depends on the incoming quality, the probability that the lot will be accepted, and the sample and lot sizes.

GATI

Creates an average total inspection (ATI) curve, which approximates the relationship between the quality of the incoming material and the number of items that need to be inspected, assuming that rejected lots will be 100% inspected and defective items will be reworked and inspected again (rectifying inspection).

The average total inspection per lot depends on the incoming quality, the probability that the lot will be accepted, and the sample and lot sizes.

GAREGION

Creates an acceptance region plot, which allows you to see the region of sample means and sample standard deviations for which you will accept a lot, when both specifications are given and the standard deviation is unknown.

ONEGRAPH

Displays all graphs on one graph.

Options

DEFECTIVES K...K

Calculates probability of acceptance for additional defective levels, specified by K...K.

WALLIS

Specify to use the approximation procedure proposed by Wallis to calculate the maximum standard deviation. This option is only available when you specify a value for both `LSPEC` and `USPEC`, but do not specify a value with `SIGMA`. If you issue `WALLIS` and the previous conditions are not met, Minitab ignores the subcommand.

Storage

SSIZE C

Stores sample size in worksheet, in column C.

SANUMBER C

Stores accept number in worksheet, in column C.

SDEFECTIVES C

Stores the defectives/defect levels in worksheet, in column C. When this command is issued, Minitab stores the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SPACCEPT C

Stores the probability of accepting values in worksheet, in column C. When this command is issued, Minitab stores the acceptance values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SPREJECT C

Stores the probability of rejecting values in worksheet, in column C. When this command is issued, Minitab stores the acceptance values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES. The probability of rejecting is 1 – the probability of accepting.

SAOQ C

Stores the average outgoing quality values in the worksheet, in column C. When this command is issued, Minitab stores the AOQ values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

SATI C

Stores the average total inspection values in the worksheet, in column C. When this command is issued, Minitab stores the ATI values associated with the AQL, RQL, and any additional defective levels specified with DEFECTIVES.

ARDECISION: Session command for accepting or rejecting an entire lot

ARDECISION

Allows you to enter actual measurements from a sample and determine whether you should accept or reject the entire lot based on sampling plan criteria.

Plan criteria

LSPEC K

Specifies the lower specification limit for the measurements. You must enter at least one specification limit.

USPEC K

Specifies the upper specification limit for the measurements. You must enter at least one specification limit.

SIGMA K

Optionally, specifies the standard deviation.

CDISTANCE K

Specifies the critical distance K, that is calculated from creating the sampling plan, and uses K to accept or reject a lot. K is a positive number.

WALLIS

Specify to use the approximation procedure proposed by Wallis to calculate the maximum standard deviation. This option is only available when you specify a value for both `LSPEC` and `USPEC`, but do not specify a value with `SIGMA`. If you issue `WALLIS` and the previous conditions are not met, Minitab ignores the subcommand.

Reliability/Survival

Test Plans

DTESTPLAN: Session command for creating a demonstration test plan

DTESTPLAN K...K

Generates a demonstration test plan with a maximum number of K failures allowed.

A demonstration test plan is to show that you have met a reliability specification or that a redesigned system has suppressed or significantly reduced the occurrence of a known cause of failure. In a demonstration test, you verify that only a certain number of failures occur in a set amount of test time.

Minimum value to be demonstrated

SCLOCATION K

Specifies the scale parameter for the Weibull or exponential distribution, or the location parameter for the other distributions.

PTILE K

Specifies the percentile for the given percent. The argument of PTILE is the percent and the argument of TIME is the percentile.

TIME K

Demonstrates that the specified percentile exceeds K. The argument of time is the percentile.

RELIABILITY K

Specifies the reliability goal as K at time K.

TIME K

Demonstrates that the reliability exceeds the specified value at time K. The argument of time is the percentile.

MTTF K

Specifies the mean time to failure as K.

Sample sizes or testing times

Specifies either sample sizes or testing times.

SAMPLE K...K

Specifies the sample size for each test plan. Minitab calculates the corresponding testing time.

TTIME K...K

Specifies the testing time for each test plan. Minitab calculates the corresponding sample sizes.

Planning distributions

Specifies the assumed distribution of failure data to be collected.

WEIBULL
Weibull distribution.

EXPONENTIAL
Exponential distribution.

SEXTREMEVALUE
Smallest extreme value distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal (base e) distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

SHSCALE K
Specifies the assumed parameter. For the Weibull distribution, K is the shape. For all other distributions, K is the scale. Do not use the SHSCALE subcommand with the exponential distribution.

Graphs

GPOPGRAPH
Displays the probability of passing a test plan graph.

OVMAXFAILURES
Use if you have requested more than one test plan (maximum number of failures) and you want to overlay the probability of passing each test plan on the same graph.

XMINIMUM K
Specify the minimum improvement ratio or amount of improvement to be displayed on the x-axis.

The x-axis is in terms of the ratio of improvement for the Weibull, lognormal, exponential, and loglogistic distributions. It is in terms of the amount of improvement for the normal, logistic, and extreme value distributions.

XMAXIMUM K
Specify the maximum improvement ratio or amount of improvement to be displayed on the x-axis.

The x-axis is in terms of the ratio of improvement for the Weibull, lognormal, exponential, and loglogistic distributions. It is in terms of the amount of improvement for the normal, logistic, and extreme value distributions.

Options

CONFIDENCE K
Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ETESTPLAN: Session command for creating an estimation test plan

ETESTPLAN
Generates an estimation test plan.

An estimation test plan indicates the number of test units that you need to estimate percentiles or reliabilities with a specified degree of precision.

Parameter to be estimated

EPTILE K
Estimates a specific percentile at K.

ERELIABILITY K
Estimates a specific reliability. K is the corresponding time.

Precisions

DUPPER is the same as DLOWER only when the CI is symmetric about the MLE. For example, any confidence interval for the K^{th} percentile in a location-scale model is symmetric about the MLE.

DLOWER K...K
The precision given as the distance between the MLE and the lower bound of the confidence interval.

DUPPER K...K
The precision given as the distance between the MLE and the upper bound of the confidence interval

Planning distributions

Specifies the assumed distribution of failure data to be collected.

WEIBULL
Weibull distribution.

EXPONENTIAL
Exponential distribution.

SEXTREMEVALUE
Smallest extreme value distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal (base e) distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

Planning values

Specify the planning values. You must provide exactly two planning values for all distributions except the exponential.

SHSCALE K
Specifies the planning value for the shape parameter (Weibull distribution) or scale parameter (other distributions). You cannot use SHSCALE with the exponential distribution.

SETS K
Specifies the true value of the known parameter. You cannot use SHSCALE and SETS simultaneously. You cannot use SETS with the exponential distribution.

SCLOCATION
Specifies the planning value for the scale parameter (Weibull or exponential distributions) or location parameter (other distributions).

PTILE K [K]
Specifies the planning values for the percentiles. Specifies the percentile for the given percent. The arguments of PTILE are the percents and the arguments of TIME are the percentile. You can provide either one or two planning values for the percentiles. If PTILE has two arguments, then you cannot use SETS or SCLOCATION.

TIME K [K]
Specifies the planning value for the percentile as K.

Right censoring

UPTIME C...C

UPTIME K...K
Use for Type I censoring on the right (Time Censoring). Censored observations are those whose failure times are greater than or equal to the UPTIME value.

- If a column has more than one row, each row value is the censoring time for each group in a multiple censoring scheme.
- If constants are specified, Minitab assumes a singly censoring scheme, where each constant is a replicate.

If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

UPFAIL C...C

UPFAIL K...K
Use for Type II censoring on the right (Failure Censoring). Specify the upper cutoff for the number of failures. Censored observations are those at the upper cutoff point and all subsequent observations.

- If a column has more than one row, each row value is the censoring percent failure for each group in a multiple censoring scheme.
- If constants are specified, Minitab assumes a singly censoring scheme, where each constant is a replicate.

If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

PRPBROUP C...C

Each column contains the expected percent of units run per group. The rows in each column must sum to 100.

The arguments are paired with the column arguments of UPTIME or UPFAIL when your data are multiply censored. In this case, if PRPGROUP is not specified, Minitab assumes an equal percent of units run in each group.

Interval censoring

NINSPECTIONS K

Specifies the number of inspections for interval-censored data. You must provide exactly one of the following subcommands. If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

You cannot use this subcommand with UPTIME, UPFAIL, or PRPGROUP.

ESITIMES K

Specifies equally spaced inspection times. K is the last inspection time.

EPITIMES K

Specifies inspection intervals containing an equal percent of units failing. K is the expected percent of failures in the entire life test.

ESILTIMES K K

Specifies equally spaced inspections in log times. The arguments are the first and last inspection times, respectively.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

TWOSIDED

Indicates that the precision corresponds to a two-sided confidence interval. By default, Minitab assumes a one-sided confidence interval.

ALTTESTPLAN: Session command for creating an accelerated life test plan

ALTTESTPLAN

Generates an accelerated life test plan.

Use an accelerated life test plan to estimate the standard error of percentiles and reliabilities when you know the number of units that are available to test, or when you want to calculate the number of test units that you need to estimate percentiles or reliabilities with a specified degree of precision.

Parameter to be estimated

EPTILE K
Specifies that you want to estimate the K^{th} percentile.

ERELIABILITY K
Specifies that you want to estimate the reliability at time K.

Sample sizes or precisions

Specifies precisions or sample sizes. You must provide exactly one of the following subcommands.

DUPPER is the same as DLOWER only when the CI is symmetric about the MLE. For example, any confidence interval for the K^{th} percentile in a location-scale model is symmetric about the MLE.

SAMPLE K...K
The available total sample size.

DLOWER K...K
The precision given as the distance between the MLE and the lower bound of the confidence interval.

DUPPER K...K
The precision given as the distance between the MLE and the upper bound of the confidence interval.

Planning distributions

Specifies the assumed distribution of failure data to be collected.

WEIBULL
Weibull distribution.

EXPONENTIAL
Exponential distribution.

SEXTREMEVALUE
Smallest extreme value distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal (base e) distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

Relationships

Specifies the relationship between failure times or log failure times and the accelerating variable.

ARRHENIUS
Arrhenius relationship.

POWER
Power relationship (also called ln relationship).

LINEAR
Linear relationship.

INVTEMP
Inverse temperature relationship.

Planning values

Specify the planning values. You must provide exactly two planning values for all distributions except the exponential.

SHSCALE K
Specifies the planning value for the shape parameter (Weibull distribution) or scale parameter (other distributions). You cannot use SHSCALE with the exponential distribution.

SETS K
Specifies the true value of the known parameter. You cannot use SHSCALE and SETS simultaneously. You cannot use SETS with the exponential distribution.

SCLOCATION
Specifies the planning value for the scale parameter (Weibull or exponential distributions) or location parameter (other distributions).

PTILE K [K]
Specifies the K^{th} percentile as the planning value The arguments of PTILE are the percents, and the arguments of TIME are the percentile. You can provide either one or two planning values for the percentiles. If PTILE has two arguments, then you cannot use SETS or SCLOCATION.

TIME K [K]
Specifies the planning value for the percentile as K.

STRESS K [K]
Specifies the stress value for the percentile as K.

Stresses

DSTRESS K
Specifies the design stress value as K.

TSTRESSES K...K

TSTRESSES C...C
Specifies the fixed test stresses as follows:

- Use constant arguments (K...K) to represent the test stress values in a single test plan.
- Use column arguments (C...C) to represent a set of test stress values for a collection of test plans. Columns must be the same length.

You must also specify either ALLOCATIONS or INCRANGES.

ALLOCATIONS K...K

ALLOCATIONS C...C

Specifies the percent allocations of units to the test stress levels. You must provide the test stresses in order from the lowest level (not necessarily the lowest value) to the highest level (not necessarily the highest value). As a result, the stress values—including the design stress value—are either increasing or decreasing.

- Use constant arguments (K...K) to represent the percent allocations in a single test plan.
- Use column arguments (C...C) to represent the percent allocations for a set of test stress values for a collection of test plans. Columns of allocations and test stresses must be the same length.

INCRANGES K

Searches for the best allocation of test units to stress levels by including prefixed ranges in the search space. K is the step length Minitab uses to search through ranges of proportions in order to optimize the proportionate allocation of units to the test stress levels. Small K values widen the search space and as a result, the optimization process becomes more computationally intensive. The default value is 0.05.

NBESTPLANS K

Specifies the number of test plans displayed in the output in addition to those specified using the ALLOCATIONS subcommand. By default, Minitab displays the three best plans.

Right censoring

UPTIME K...K

Use for Type I censoring on the right (Time Censoring). Censored observations are those whose failure times are greater than or equal to the UPTIME value. Each argument of UPTIME represents the censoring times associated with a stress level. The first argument corresponds to the censoring time at the lowest test stress level, the second argument corresponds to the censoring time at the second test stress level, and so on.

If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

UPFAIL K...K

Use for Type II censoring on the right (Failure Censoring). Specify the upper cutoff for the percent failures. Censored observations are those at the upper cutoff point and all subsequent observations. The arguments of UPFAIL represent the percent of failures corresponding to each stress level. The first argument corresponds to percent failures at the lowest test stress level, the second argument corresponds to percent failures at the second test stress level, and so on.

If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

Interval censoring

NINSPECTIONS K

Specifies the number of inspections for interval-censored data. You must provide exactly one of the following subcommands. If none of the censoring commands (UPTIME, UPFAIL, or NINSPECTIONS) are specified, the data to be collected is assumed to be uncensored.

You cannot use this subcommand with UPTIME, UPFAIL, or PRPGROUP.

ESITIMES K

Specifies equally spaced inspection times. K is the last inspection time for each test stress level. The first argument is the last inspection time at the lowest test stress level, the second argument is the last inspection at the second test stress level, and so on.

EPITIMES K

Specifies inspection intervals containing an equal percent of units failing. K is the expected percent of failures for each test stress level. The first argument is the percent failures at the lowest test stress level, the second argument is the percent failures at the second test stress level, and so on.

ESILTIMES K K

Specifies equally spaced inspections in log times. The two arguments are the first and last inspection times, respectively for each test stress level. ESILTIMES must have exactly twice as many arguments as there are test stress levels. The first set of arguments are the first inspection time at the lowest test stress level, the first inspection time at the second test stress level, and so on. The second set of arguments are the last inspection time at the lowest test stress level, the last inspection time at the second test stress level, and so on.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

TWOSIDED

Indicates that the precision corresponds to a two-sided confidence interval. By default, Minitab assumes a one-sided confidence interval.

Distribution Analysis (Right Censoring)

RDIDENTIFICATION: Session command for creating a distribution ID plot

RDIDENTIFICATION C...C

Generates probability plots for right-censored/uncensored, failure (or survival) data.

Use RDIDENTIFICATION to plot up to eleven different probability plots (with distributions chosen from smallest extreme value, Weibull, three-parameter Weibull, exponential, two-parameter exponential, normal, lognormal, three-parameter lognormal, logistic, loglogistic, and three-parameter loglogistic) to help you determine which of these distributions best fits your data. Usually this is done by comparing how closely the plot points lie to the best-fit lines—in particular, points in the tails of the distribution.

The sample data are the individual failure times. For example, you may have collected samples of failure times under different temperatures, or under varying conditions of any combination of stress variables. The data can be singly censored on the right, with censoring of either type I (time censoring) of type II (failure censoring). You can also supply censor columns for multiple censored data, where a 0 denotes a censored observation and a 1 denotes a failure time.

You can have up to 50 input columns, or samples.

FREQUENCY C...C

Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times.

BY C

For stacked data, C is the variable that defines the groups. Use one By variable to represent your groups.

Censoring options

CENSOR C...C

If you have multiple right censored data, CENSOR specifies which observations are censored.

Use a single column for all samples or specify a column for each sample. If you specify a column for each sample, the same column can apply to multiple samples. Use two distinct values in each censoring column, C...C, as indicator variables for censored observations. By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value. You can specify which value is the censoring value by using the CVALUE subcommand. The censored observations are assumed to be right censored while the uncensored observations are considered exact failure times. Specify the columns in the same order as the samples for the main command.

CVALUE K or K...K

This subcommand allows you to specify the censoring value, **K**, which is the value in the censoring column that defines a censored observation. Specify a single value for all censoring columns or specify a value for every censoring column. When you specify multiple values, follow the order for the censoring columns with CENSOR.

UPTIME K or K...K

If you have Type I censoring on the right (time censoring), UPTIME specifies the upper cutoff for failure time. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column.

Observations whose failure times are greater than or equal to the upper cutoff are considered to be censored, observations whose failure times are not greater than the upper cutoff are uncensored. For example, UPTIME 500 says to make all observations whose failure times are >= 500 censored, all observations whose failure times are < 500 uncensored. Specify the values in the same order as the samples for the main command.

UPFAIL K or K...K

If you have Type II censoring on the right (failure censoring), UPFAIL specifies the upper cutoff for the number of failures. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column. Specify the values in the same order as the samples for the main command.

The observation at the upper cutoff point, and all subsequent observations, are considered to be censored. Observations prior to the upper cutoff point are uncensored. For example, UPFAIL 20 says to make all (ordered) observations starting with the 20th observations censored, all (ordered) observations prior to the 20th observation uncensored.

Estimation methods

PTILES K...K

The percents K that define the additional percentiles to be estimated.

LSXY

Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE
Use MLE to estimate parameters using maximum likelihood method.

Distributions

Use the following subcommands to create distribution plots of the following distributions.

SEXTREMEVALUE
Smallest extreme value distribution.

WEIBULL
Weibull distribution.

TWEIBULL
Three-parameter Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
Two-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal distribution.

TLNORMAL
Three-parameter lognormal distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Loglogistic distribution.

TLLOGISTIC
Three-parameter loglogistic distribution.

Probability plot options

Use one of the following subcommands to specify how plot points are determined. Minitab uses the median rank (Benard) method as the default, if you do not specify one of these subcommands.

MEDRANK
Median ranks (Benard) method.

HJ
Herd-Johnson method.

MKM
Modified Kaplan-Meier method.

KM
Kaplan-Meier method.

ALLPTS
For identical failure times, ALLPTS plots the cumulative percents for each occurrence of a failure time.

AVEPTS
For identical failure times, AVEPTS plots the median cumulative percent of identical failure times.

MAXPTS
For identical failure times, MAXPTS plots the maximum cumulative percent of identical failure times.

Graph options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

SEPARATE
Displays each graph of different variables or By variable in a separate graph. Minitab overlays graphs, if you omit this subcommand.

XMINIMUM K
For the endpoints of the x-axis, XMINIMUM specifies the minimum value, K, along the x-axis.

XMAXIMUM K
For the endpoints of the x-axis, XMAXIMUM specifies the maximum value, K, along the x-axis.

RDOVIEW: Session command for creating a layout of distribution plots

RDOVIEW C...C
Generates distribution plots of specified distributions for right-censored, failure time data, specified in columns C...C.
RDOVIEW generates a layout of distribution plots for data for up to 50 samples of individual failure times. For example, you may have collected samples of failure times under different temperatures, or under varying conditions of any combination of stress variables. The data can be singly censored on the right, with censoring of either type I (time censoring) of type II (failure censoring). You can also supply censor columns for multiple censored data, where a 0 denotes a censored observation and a 1 denotes a failure time.

If you choose a parametric analysis, Minitab displays four different ways of describing the distribution of failure time data in each sample - a probability density function plot, a probability plot, a survival plot, and a hazard plot. The probability plot displays (usually transformed) estimates of the cumulative distribution function F(y) versus (usually transformed) failure time. The survival plot displays the survival or reliability function 1-F(y) vs. failure time. The hazard plot displays the hazard function f(y)/(1-F(y)) versus failure time.

If you choose a nonparametric analysis, Minitab displays a survival plot and a hazard plot.

FREQUENCY C...C
Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times.

BY C

For stacked data, C is the variable that defines the groups. Use one By variable to represent your groups.

Censoring options

CENSOR C...C

If you have multiple right censored data, CENSOR specifies which observations are censored.

Use a single column for all samples or specify a column for each sample. If you specify a column for each sample, the same column can apply to multiple samples. Use two distinct values in each censoring column, C...C, as indicator variables for censored observations. By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value. You can specify which value is the censoring value by using the CVALUE subcommand. The censored observations are assumed to be right censored while the uncensored observations are considered exact failure times. Specify the columns in the same order as the samples for the main command.

CVALUE K or K...K

This subcommand allows you to specify the censoring value, **K**, which is the value in the censoring column that defines a censored observation. Specify a single value for all censoring columns or specify a value for every censoring column. When you specify multiple values, follow the order for the censoring columns with CENSOR.

UPTIME K or K...K

If you have Type I censoring on the right (time censoring), UPTIME specifies the upper cutoff for failure time. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column.

Observations whose failure times are greater than or equal to the upper cutoff are considered to be censored, observations whose failure times are not greater than the upper cutoff are uncensored. For example, UPTIME 500 says to make all observations whose failure times are >= 500 censored, all observations whose failure times are < 500 uncensored. Specify the values in the same order as the samples for the main command.

UPFAIL K or K...K

If you have Type II censoring on the right (failure censoring), UPFAIL specifies the upper cutoff for the number of failures. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column. Specify the values in the same order as the samples for the main command.

The observation at the upper cutoff point, and all subsequent observations, are considered to be censored. Observations prior to the upper cutoff point are uncensored. For example, UPFAIL 20 says to make all (ordered) observations starting with the 20th observations censored, all (ordered) observations prior to the 20th observation uncensored.

Parametric estimation methods

LSXY

Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE

Use MLE to estimate parameters using maximum likelihood method.

Nonparametric estimation methods

KMESTIMATE
Specifies the Kaplan-Meier method for nonparametric estimates.

INTBY K K
Specifies the intervals for the actuarial method. The intervals are created from 0 to K by K.

Distributions

Use the following subcommands to create distribution plots of the following distributions.

SEXTREMEVALUE
Smallest extreme value distribution.

WEIBULL
Weibull distribution.

TWEIBULL
Three-parameter Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
Two-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal distribution.

TLNORMAL
Three-parameter lognormal distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Loglogistic distribution.

TLLOGISTIC
Three-parameter loglogistic distribution.

Probability plot options

Use one of the following subcommands to specify how plot points are determined. Minitab uses the median rank (Benard) method as the default, if you do not specify one of these subcommands.

MEDRANK
Median ranks (Benard) method.

HJ
Herd-Johnson method.

MKM
Modified Kaplan-Meier method.

KM
Kaplan-Meier method.

ALLPTS
For identical failure times, ALLPTS plots the cumulative percents for each occurrence of a failure time.

AVEPTS
For identical failure times, AVEPTS plots the median cumulative percent of identical failure times.

MAXPTS
For identical failure times, MAXPTS plots the maximum cumulative percent of identical failure times.

Graph options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

SEPARATE
Displays each graph of different variables or By variable in a separate graph. Minitab overlays graphs, if you omit this subcommand.

XMINIMUM
For the endpoints of the x-axis, XMINIMUM specifies the minimum value, K, along the x-axis.

XMAXIMUM
For the endpoints of the x-axis, XMAXIMUM specifies the maximum value, K, along the x-axis.

LTEST: Session command for parametric or nonparametric distribution analysis

LTEST C...C
LTEST fits a smallest extreme value, Weibull, three-parameter Weibull, exponential, two-parameter exponential, normal, lognormal, three-parameter lognormal, logistic, loglogistic, or three-parameter loglogistic distribution to the data that you specify in columns C...C.

From the distribution that you chose, estimate percentiles, survival probabilities, cumulative failure probabilities, evaluate the appropriateness of the chosen distribution, make survival, cumulative failure, hazard and probability plots, and so on. Your data can be right censored or exact failure times. You can have multiple failure modes.

If no distribution fits your data, then Minitab provides nonparametric estimates of survival probabilities, hazard function, and so on.

NOPARAMETRIC
When no distribution fits your data, Minitab provides nonparametric estimates of survival probabilities, hazard function, and so on.

Nonparametric estimation methods are Kaplan-Meier and Actuarial methods.

FREQUENCY C...C

Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times.

BY C

For stacked data, **C** is the variable that defines the groups. Use one By variable to represent your groups.

Censoring options

CENSOR C...C or C...C

If you have multiple right censored data, CENSOR specifies which observations are censored.

Use a single column for all samples or specify a column for each sample. If you specify a column for each sample, the same column can apply to multiple samples. Use two distinct values in each censoring column, C...C, as indicator variables for censored observations. By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value. You can specify which value is the censoring value by using the CVALUE subcommand. The censored observations are assumed to be right censored while the uncensored observations are considered exact failure times. Specify the columns in the same order as the samples for the main command.

CVALUE K or K...K

This subcommand allows you to specify the censoring value, **K**, which is the value in the censoring column that defines a censored observation. Specify a single value for all censoring columns or specify a value for every censoring column. When you specify multiple values, follow the order for the censoring columns with CENSOR.

UPTIME K or K...K

If you have Type I censoring on the right (time censoring), UPTIME specifies the upper cutoff for failure time. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column.

Observations whose failure times are greater than or equal to the upper cutoff are considered to be censored, observations whose failure times are not greater than the upper cutoff are uncensored. For example, UPTIME 500 says to make all observations whose failure times are >= 500 censored, all observations whose failure times are < 500 uncensored. Specify the values in the same order as the samples for the main command.

UPFAIL K or K...K

If you have Type II censoring on the right (failure censoring), UPFAIL specifies the upper cutoff for the number of failures. This command applies only to single censoring. Specify a single value for all sample columns or specify a value for every sample column. Specify the values in the same order as the samples for the main command.

The observation at the upper cutoff point, and all subsequent observations, are considered to be censored. Observations prior to the upper cutoff point are uncensored. For example, UPFAIL 20 says to make all (ordered) observations starting with the 20th observations censored, all (ordered) observations prior to the 20th observation uncensored.

Distributions

Use these subcommands to analyze data assuming one of these distributions.

SEXTREMEVALUE

Smallest extreme value distribution.

WEIBULL

Weibull distribution.

TWEIBULL
Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
2-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal (base e) distribution.

TLNORMAL
Lognormal (base e) distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

Failure modes

FMODE c or C...C
Estimates overall system reliability and component reliability when multiple causes of failure are present. Specify a single column for all samples or a column for every sample. When you specify a column for every sample, repeat columns are valid. Specify the columns in the same order as the samples for the main command.

KEEP K...K
Specifies which failure modes to include in the analyses.

There may be times when you are only interested in analyzing some of the possible failure modes. For example, you may wish to analyze the overall system reliability with only 1 of the 3 possible failure modes. Specify levels from the column of failure modes to include only those modes in the analysis.

- If the column is numeric, then specify numbers that represent the failure modes.
- If the column is text, then specify text labels that represent failure modes.

ELIMINATE K...K
Specifies which failure modes to exclude from the analyses.

There may be times when you are only interested in analyzing some of the possible failure modes. For example, you may wish to analyze the overall system reliability with only 2 of the 3 possible failure modes. Specify levels from the column of failure modes to exclude those modes from the analysis

- If the column is numeric, then specify numbers that represent the failure modes.
- If the column is text, then specify text labels that represent failure modes.

CDISTRIBUTION K *"text"*...K *"text"*
Specifies which failure modes to change the distribution of, and which distribution to use. For example, to change the distribution of failure mode A to exponential, type "A" "EXPONENTIAL".

Change the distribution of a failure mode from the assumed overall distribution when a particular failure mode is well modeled by a specific distribution. For example, you specify LTEST; WEIBULL, for the overall system, but change the distribution for failure mode 1 to be lognormal and for failure mode 2 to be normal.

Choose from the distributions described above.

Use K to specify the level of a failure mode that you want to change the distribution for, then type the distribution in double quotes. For example `CDISTRIBUTION "Obstruction" "EXPONENTIAL"` specifies to use the exponential distribution for the failure mode with the label *Obstruction*.

RINDANALYSES
Displays results for individual failure modes.

SINDANALYSES
Stores results for individual failure modes.

Parametric estimation methods

Choose either the LSXY or MLE parametric estimation method.

LSXY
Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE
Use MLE to estimate parameters using maximum likelihood method.

COMSHAPE
Use COMSHAPE to assume common shape parameter.

COMSCALE
Use COMSCALE to assume common scale parameter.

TIMES K...K

TIMES C
Calculate the survival probabilities or cumulative failure probabilities for times K...K or C

ECFPROBABILITY
Choose to estimate cumulative failure probabilities instead of survival probabilities.

STIME
Stores the times associated with the survival probabilities or cumulative failure probabilities.

SURVPROBABILITY C...C
Stores the survival probability for **T**. Where **T** is given on the TIME subcommand.

CLSURVPROBABILITY C...C
Stores the upper and lower confidence limits of the survival probability.

CFPROBABILITY C...C
Stores the cumulative failure probability for **T**. Where **T** is given on the TIME subcommand.

CLCFPROBABILITY C...C
Stores the upper and lower confidence limits of the cumulative failure probability.

PTILES K...K

PTILES C

The percents K which define the additional percentiles to be estimated.

You can enter individual percents (0 < P < 100) or a column of percents.

Bayes analysis

Choose to set the shape (SETSHAPE) or scale (SETSCALE) parameters depending on your distribution.

- For the Weibull distribution, estimate the scale parameter while keeping the shape parameter fixed. The number of arguments must be 1 or equal to the number of variables. If there is 1 argument then all distributions will have the same shape parameter.
- For all distributions except the Weibull distribution, estimate the location parameter while keeping the scale parameter fixed. The number of arguments must be 1 or equal to the number of variables. If there is 1 argument then all distributions will have the same scale parameter.

SETSHAPE K...K

Sets the value for the shape parameter of the Weibull distribution. Enter a list of values, K...K, for the shape that is equal in length to the number of samples.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

If the analysis mixes the Weibull distribution with location-scale distributions, then specify all the shape and scale parameters with the same subcommand. Use the order from the failure modes and samples the same as if all the distributions were the same.

SETSCALE K...K

Sets the value for the scale parameter. Enter a list of values, K...K, for the scale that is equal in length to the number of samples.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First

Order	Failure mode	Sample
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

If the analysis mixes the Weibull distribution with location-scale distributions, then specify all the shape and scale parameters with the same subcommand. Use the order from the failure modes and samples the same as if all the distributions were the same.

SETTHRESHOLD K...K

Sets the value for the threshold parameter. Enter a list of values, K...K, for the threshold that is equal in length to the number of samples. To set threshold parameters, all the distributions must have a threshold parameter.

For all distributions with threshold parameters (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic) you can fix the threshold parameter.

The threshold parameter you set must be less than the minimum data value.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

Nonparametric estimation methods

Use KMESTIMATE, INTBY, or INTENDPOINTS as the nonparametric estimation method.

KMESTIMATE

Use the Kaplan-Meier method for nonparametric estimates.

INTBY K K

Specify the intervals for the actuarial method. The intervals are created from 0 to K by K.

INTENDPOINTS K...K

Specify the endpoints of the intervals for the actuarial method. The lower endpoint of the first interval is given by the first K. The next value is the end of the first interval and the beginning of the second interval. The subsequent values are also beginning and ends of intervals until the last K. The last K is the endpoint of the last complete interval. If data points are, greater than the last K, then the analysis also makes a right-censored interval for those points.

ECFPROBABILITY

Choose to estimate cumulative failure probabilities instead of survival probabilities.

STIME C
Stores the times associated with the nonparametric survival probabilities or cumulative failure probabilities.

SURVPROBABILITY C...C
Stores the nonparametric survival probability for **T**. Where **T** is given on the TIME subcommand.

SESURVPROBABILITY C...C
Stores the standard error of the nonparametric survival probability.

CLSURVPROBABILITY C...C
Stores the upper and lower confidence limits of the nonparametric survival probability.

CFPROBABILITY C...C
Stores the nonparametric cumulative failure probability for **T**. Where **T** is given on the TIME subcommand.

SECFPROBABILITY C...C
Stores the standard error of the nonparametric cumulative failure probability.

CLCFPROBABILITY C...C
Stores the upper and lower confidence limits of the nonparametric cumulative failure probability.

HAZARD C...C
Stores hazard estimates.

HTIME C...C
Stores times associated with hazard estimates.

Confidence intervals

CONFIDENCE K
Specifies the percent confidence, K, for your confidence intervals.

LOWCI
Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI
Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI
Calculates two-sided confidence intervals for parameter estimates and percentiles.

Parameter tests

TVSS K
Tests whether the shape (Weibull) or scale (other distributions) is equal to K (K>0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the shape parameter equals 2 for two different failure modes, specify `TVSS 2 2;`

TVSL K
Tests whether the scale (Weibull or exponential) or location (other distributions) is equal to K (K>0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the scale parameter equals 2 for two different failure modes, specify `TVSL 2 2;`

TVTHRESHOLD K

Tests whether the threshold (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic) is equal to K (K>0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the threshold parameter equals 2 for two different failure modes, specify `TVTHRESHOLD 2 2;`

TESS

Tests for equal shape (Weibull) or scale (other distributions).

TESL

Tests for equal scale (Weibull or exponential) or location (other distributions).

TETHRESHOLD

Tests for equal threshold (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic).

Graph options

CI

Displays confidence intervals on probability and survival plots.

PPLOT

Displays a probability plot.

DCPPLOT

Specifies to include symbols for right-censored and left-censored observations on the probability plot.

MEDRANK

MEDRANK specifies the median ranks (Benard) method for the plotted points on the probability plot. Minitab uses the median rank (Benard) method as the default, if you do not specify MEDRANK, HJ, MKM, or KM.

HJ

HJ specifies the Herd-Johnson method for the plotted points on the probability plot.

MKM

MKM specifies the modified Kaplan-Meier method for the plotted points on the probability plot.

KM

KM specifies the Kaplan-Meier method for the plotted points on the probability plot.

ALLPTS

For identical failure times, ALLPTS plots the cumulative percents for each occurrence of a failure time.

AVEPTS

For identical failure times, AVEPTS plots the median cumulative percent of identical failure times.

MAXPTS

For identical failure times, MAXPTS plots the maximum cumulative percent of identical failure times.

SPLOT

Displays a parametric or nonparametric survival plot. The nonparametric survival plot is a step function, while the parametric survival plot is a smooth curve.

CFPLOT

Displays a parametric or nonparametric cumulative failure plot. The nonparametric cumulative failure plot is a step function, while the parametric cumulative failure plot is a smooth curve.

HPLOT

Displays a parametric or nonparametric hazard plot. The nonparametric hazard plot is a step function, while the parametric hazard plot is a smooth curve.

PANEL

Use PANEL and SEPARATE to perform hypothesis tests on the distribution parameters.

PANEL displays each graph of different variables or By variable in a separate panel of the graph.

SEPARATE

Use PANEL and SEPARATE to perform hypothesis tests on the distribution parameters.

SEPARATE displays each graph of different variables or By variable in a separate graph.

XMINIMUM K

For parametric distribution analysis, specify the minimum and maximum X scale for the probability, survival and hazard plots. For nonparametric distribution analysis, specify the minimum and maximum X scale for the survival and hazard plot.

Use XMINIMUM to specify the X-axis scale minimum.

XMAXIMUM K

For parametric distribution analysis, specify the minimum and maximum X scale for the probability, survival and hazard plots. For nonparametric distribution analysis, specify the minimum and maximum X scale for the survival and hazard plot.

Use XMAXIMUM to specify the X-axis scale maximum.

XLABEL K

XLABEL *"text"*

Use XLABEL to specify an x-axis label for the graph. When you omit this subcommand, Minitab displays a default label.

Results for parametric and nonparametric analyses

NODEFAULTS

Specifies that no default tables and graphs will be displayed.

TMETHOD

Displays the method table. The table includes information about the estimation method, the number of missing rows, frequency, and any failure modes.

TCENSOR

Displays the censoring table. The table provides information about the censoring scheme, such as any censoring value and the numbers of rows with each type of censoring.

TCDISTRIBUTION

For a parametric analysis, displays the characteristics of the distribution table. The table includes point estimates, standard errors, and confidence intervals for the characteristics of the sample distribution. The characteristics include the Mean Time To Failure (MTTF), standard deviation, median, first quartile, third quartile, and interquartile range. For the least squares (failure time(X) on rank(Y)) estimation method the table provides only point estimates.

For a non-parametric analysis, displays the characteristics of the variable table. The content of the table depends on the estimation method. For the actuarial method of estimation, the table displays point estimates, standard errors, and confidence intervals for the median time to failure. For the Kaplan-Meier estimation method, the table displays these estimates for the mean. For the Kaplan-Meier estimation method, the table also includes point estimates of the first quartile, third quartile, and interquartile range.

TPROBABILITY

Displays a table of survival or cumulative failure probabilities. The table includes estimates and confidence intervals of survival or cumulative failure probabilities.

For a parametric analysis, the probabilities are for the values from the `TIMES` subcommand. Without the `TIMES` subcommand, the table does not display.

For the least squares (failure time(X) on rank(Y)) estimation method the table provides only point estimates of probabilities.

The title of the table depends on the estimation method. For the estimation method is Kaplan-Meier, the title uses the estimation method. For other methods, the title uses the type of probabilities.

Results for a parametric analysis

STEP

Displays the iteration history table. The table includes the log-likelihood at each iterative step of the optimization algorithm that estimates the distribution parameters.

TPARAMETER

Displays the parameter estimates table. The table includes point estimates for distribution parameters. For the maximum likelihood estimation method, the table also includes standard errors and confidence intervals.

TGOODNESS

Displays the goodness-of-fit table. For maximum likelihood estimation, the table includes the adjusted Anderson-Darling statistic. For least squares estimation the table includes the adjusted Anderson-Darling statistic and the correlation coefficient.

TPERCENTILE

Displays the table of percentiles. The table includes the estimated percentiles of the distribution. For the maximum likelihood estimation method, the table also includes standard errors and confidence intervals of the percentiles.

TTESTS

Displays the results of the tests of the parameters. The tests are present when the analysis includes the subcommand for that test.

Results for a nonparametric analysis

TATIMES

Displays a table of additional time from time T until 50% of running units fail. The table includes point-estimates of additional survival time from the interval start times. The table also includes standard errors and confidence intervals. The table is available with the actuarial estimation method.

TACTUARIAL

Displays a table of actuarial estimates. The table is available with the acturial estimation method. The table includes estimated conditional failure probabilities and standard errors associated with the intervals that the INTBY or INTENDPOINTS specify.

THAZARD

Displays the hazards and densities table. With the actuarial estimation method, the table includes values of the nonparametric hazard function and the nonparametric density function. For right censored or uncensored data with the Kaplan-Meier estimation method, the table displays the empirical hazard function values.

TLRANK

Displays the log-rank statistic table for comparing multiple survival curves. The table is available for right-censored or uncensored data. The data must also have multiple samples.

TVCLRANK

Displays the variance-covariance matrix of the log-rank statistics for comparing multiple survival curves. The table is available for right-censored or uncensored data. The data must also have multiple samples.

TWILCOXON

Displays a Wilcoxon statistic table for comparing multiple survival curves. The table is available for right-censored or uncensored data. The data must also have multiple samples.

TVCWILCOXON

Displays the variance-covariance matrix of the Wilcoxon statistics for comparing multiple survival curves. The table is available for right-censored or uncensored data. The data must also have multiple samples.

TCHISQTESTS

Displays the test statistics table. The table shows the chi-square test statistics and p-values for the log-rank and Wilcoxon tests. The table is available for right-censored or uncensored data. The data must also have multiple samples.

Estimate parameters

START C...C

Specifies the initial values for the estimated parameters. If one column is given then the starting estimates apply for all variables. Otherwise, a column must correspond to each variable.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

ITERATION K
Changes the maximum number of iterations that Minitab will perform to reach convergence. The default value is 20. LTEST obtains maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

Storage of fitted distributions

PPERCENTS C
Stores the percents associated with the percentiles.

PERCENTILE C...C
Stores the percentiles.

SEPERCENTILE C...C
Stores the standard error of the percentiles.

CLPERCENTILE C...C
Stores the upper and lower confidence limits of the percentiles.

Storage of estimates

PARAM C...C
Stores the estimated parameters for the distribution down the column.

SSEPARAM C...C
Stores the standard errors of the estimated parameters for the distribution down a column.

CLPARAM C...C
Stores the confidence limits for the estimated parameters for the distribution down a column.

VCPARAM *M...M*
Stores the variance covariance matrix for the estimated parameters.

LOGLIKELIHOOD K...K
Stores the last log-likelihood.

Subcommand

CTOLERANCE K K
Changes the convergence criteria. Both Ks must be positive numbers.

For the first K, convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than **K**. The default is 0.0001.

For the second K, convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than **K**. The default is 0.0001.

Distribution Analysis (Interval Censoring)

IDIDENTIFICATION: Session command for creating probability plots of interval-censored, failure (or survival) data

IDIDENTIFICATION C...C

Use IDIDENTIFICATION to create up to eleven different probability plots (with distributions chosen from smallest extreme value, Weibull, three-parameter Weibull, exponential, two-parameter exponential, normal, lognormal, three-parameter lognormal, logistic, loglogistic, and three-parameter loglogistic) to help you determine which of these distributions best fits your data. Usually, this is done by comparing how closely the plot points lie to the best-fit lines—in particular, points in the tails of the distribution.

The failure times can be interval-censored data or warranty data.

Interval-censored data

Your data may be right, left, or interval censored, in addition to exact failure times. You may have multiple failure modes. You can have up to 50 input columns. For multiple samples in the same columns, you can have up to 50 samples.

The first K/2 columns represent start times and the remaining columns represent end times, where K is the number of columns on the command line. Therefore, LTABLE will perform K/2 analyses. Each analysis provides identical results to the results you would get if you ran each set of variables separately.

Observation	Variables	Meaning
Exact failure time	Start = End = TIME	Exact failure time = TIME
Right censored	Start = TIME, End = *	Right censored time = TIME
Left censored	Start = *, End = TIME	Left censored time = TIME
Interval censored	Start = TIME1, end = TIME2	Interval censored = (TIME1, TIME2)

Warranty data

To use warranty data, specify the WDATA subcommand after the main command. The data have two valid arrangements. In both arrangements, warranty returns are in a triangular matrix. The quantities of units that ship are in either their own column or in the first row of the columns that contain the warranty returns. Use the SHPROW subcommand to specify that the quantities are in the first row of the columns. Each arrangement allows only 1 sample at a time. Use the STWDATA subcommand to store the warranty data as interval-censored data to use with the WPREDICTIONS command.

Shipment quantities are in a column

Enter the column of shipment quantities first after IDIDENTIFICATION. Enter the remaining columns in chronological order from earliest to latest. For example, for the following data, enter

```
IDIDENTIFICATION C2-C6;
WDATA;
```

In this worksheet, the number of items shipped in each year is recorded in a single column, with each row representing a different year. The number of items returned each year is recorded in a separate column. For example, in 2011, 15,000 items were shipped. 78 of those items were returned in 2012, 130 of those items were returned in 2013, and 200 of those items were returned in 2014.

C1	C2	C3	C4	C5	C6
Year Shipped	Quantity Shipped	Returned Year 1	Returned Year 2	Returned Year 3	Returned Year 4
2010	10000	50	60	100	300
2011	15000		78	130	200
2012	23000			275	600
2013	32000				400

Shipment quantities are in the first row of the return columns

Enter all the columns in chronological order. For example, for the following data, enter

```
LTABLE C1-C4;
WDATA;
SHPROW;
```

In this worksheet, the first row indicates the total number of items shipped in each year. Each column represents a different year of shipment. Each cell of a column indicates the number of items returned in each subsequent time period. For example, 23,000 items were shipped in 2012. In the following year, 275 of those items were returned. In the year after that, 600 were returned.

C1	C2	C3	C4
Shipped Year 1	Shipped Year 2	Shipped Year 3	Shipped Year 4
10000	15000	23000	32000
50	78	275	400
60	130	600	
100	200		
300			

WDATA

Specifies that the columns arguments for the main command contain warranty data.

SHPROW

Indicates that you entered shipment values in the first row of each failure/return column. Without this subcommand, Minitab assumes the shipment values reside in a separate column.

FREQUENCY C...C

Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times. Use this subcommand only with interval-censored data.

BY C

For stacked data, C is the variable that defines the groups. Use one By variable to represent your groups. Use this subcommand only with interval-censored data.

Estimation methods

PTILES K...K

The percents K that define the additional percentiles to be estimated.

LSXY
Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE
Use MLE to estimate parameters using maximum likelihood method.

Distributions

Use the following subcommands to create distribution plots of the following distributions.

SEXTREMEVALUE
Smallest extreme value distribution.

WEIBULL
Weibull distribution.

TWEIBULL
Three-parameter Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
Two-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal distribution.

TLNORMAL
Three-parameter lognormal distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Loglogistic distribution.

TLLOGISTIC
Three-parameter loglogistic distribution.

Graph options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

SEPARATE
Displays each graph of different variables or By variable in a separate graph. Minitab overlays graphs, if you omit this subcommand.

XMINIMUM K
For the endpoints of the x-axis, XMINIMUM specifies the minimum value, K, along the x-axis.

XMAXIMUM K

For the endpoints of the x-axis, XMAXIMUM specifies the maximum value, K, along the x-axis.

Storage of data

STWDATA

Stores the transformed warranty data in 3 columns after the last column in the worksheet. The first column stores the start times, the second column stores the end times, and the last column stores the frequencies. These columns, containing interval censored data, are the input columns for the Warranty Prediction command.

IDOVIEW: Session command for creating a layout of distribution plots of interval-censored data

IDOVIEW C...C

IDOVIEW generates a layout of distribution plots. For example, you may have collected samples of failure times under different temperatures, or under varying conditions of any combination of stress variables.

If you choose a parametric analysis, Minitab displays four different ways of describing the distribution of failure time data in each sample—a probability density function plot, a probability plot, a survival plot, and a hazard plot.

If you choose a nonparametric analysis, Minitab displays a survival plot for the Turnbull method and a survival plot and a hazard plot for the actuarial method.

The failure times can be interval-censored data or warranty data.

Interval-censored data

Your data may be right, left, or interval censored, in addition to exact failure times. You may have multiple failure modes. You can have up to 50 input columns. For multiple samples in the same columns, you can have up to 50 samples.

The first K/2 columns represent start times and the remaining columns represent end times, where K is the number of columns on the command line. Therefore, `LTABLE` will perform K/2 analyses. Each analysis provides identical results to the results you would get if you ran each set of variables separately.

Observation	Variables	Meaning
Exact failure time	Start = End = TIME	Exact failure time = TIME
Right censored	Start = TIME, End = *	Right censored time = TIME
Left censored	Start = *, End = TIME	Left censored time = TIME
Interval censored	Start = TIME1, end = TIME2	Interval censored = (TIME1, TIME2)

Warranty data

To use warranty data, specify the `WDATA` subcommand after the main command. The data have two valid arrangements. In both arrangements, warranty returns are in a triangular matrix. The quantities of units that ship are in either their own column or in the first row of the columns that contain the warranty returns. Each arrangement allows only 1 sample at a time. Use the `STWDATA` subcommand to store the warranty data as interval-censored data to use with the `WPREDICTIONS` command.

Shipment quantities are in a column

Enter the column of shipment quantities first after `IDOVIEW`. Enter the remaining columns in chronological order from earliest to latest. For example, for the following data, enter

```
IDOVIEW C2-C6;
WDATA;
```

In this worksheet, the number of items shipped in each year is recorded in a single column, with each row representing a different year. The number of items returned each year is recorded in a separate column. For example, in 2011, 15,000 items were shipped. 78 of those items were returned in 2012, 130 of those items were returned in 2013, and 200 of those items were returned in 2014.

C1	C2	C3	C4	C5	C6
Year Shipped	Quantity Shipped	Returned 2011	Returned 2012	Returned 2013	Returned 2014
2010	10000	50	60	100	300
2011	15000		78	130	200
2012	23000			275	600
2013	32000				400

Shipment quantities are in the first row of the return columns

Enter all the columns in chronological order. For example, for the following data, enter

```
LTABLE C1-C4;
WDATA;
SHPROW;
```

In this worksheet, the first row indicates the total number of items shipped in each year. Each column represents a different year of shipment. Each cell of a column indicates the number of items returned in each subsequent time period. For example, 23,000 items were shipped in 2012. In the following year, 275 of those items were returned. In the year after that, 600 were returned.

C1	C2	C3	C4
Shipped 2010	Shipped 2011	Shipped 2012	Shipped 2013
10000	15000	23000	32000
50	78	275	400
60	130	600	
100	200		
300			

WDATA

Specifies that the columns arguments for the main command contain warranty data.

SHPROW

Indicates that you entered shipment values in the first row of each failure/return column. Without this subcommand, Minitab assumes the shipment values reside in a separate column.

FREQUENCY C...C

Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times. Use this subcommand only with interval-censored data.

BY C
For stacked, interval-censored data, C is the variable that defines the groups. Use one By variable to represent your groups. Use this subcommand only with interval-censored data.

Parametric estimation methods

LSXY
Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE
Use MLE to estimate parameters using maximum likelihood method.

Nonparametric estimation methods

TURNBULL
Displays a Turnbull survival plot.

ACTURIAL
Displays an actuarial survival plot and hazard plot.

Distributions

Use the following subcommands to create distribution plots of the following distributions.

SEXTREMEVALUE
Smallest extreme value distribution.

WEIBULL
Weibull distribution.

TWEIBULL
Three-parameter Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
Two-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Lognormal distribution.

TLNORMAL
Three-parameter lognormal distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Loglogistic distribution.

TLLOGISTIC
Three-parameter loglogistic distribution.

Graph options

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

SEPARATE
Displays each graph of different variables or By variable in a separate graph. Minitab overlays graphs, if you omit this subcommand.

XMINIMUM K
For the endpoints of the x-axis, XMINIMUM specifies the minimum value, K, along the x-axis.

XMAXIMUM K
For the endpoints of the x-axis, XMAXIMUM specifies the maximum value, K, along the x-axis.

Storage of data

STWDATA
Stores the transformed warranty data in 3 columns after the last column in the worksheet. The first column stores the start times, the second column stores the end times, and the last column stores the frequencies. These columns, containing interval censored data, are the input columns for the Warranty Prediction command.

LTABLE: Session command for fitting a distribution to interval-censored data or warranty data

LTABLE C...C
LTABLE fits a smallest extreme value, Weibull, three-parameter Weibull, exponential, two-parameter exponential, normal, lognormal, three-parameter lognormal, logistic, loglogistic, or three-parameter loglogistic distribution to your data. From the distribution you chose, estimate percentiles, survival probabilities, cumulative failure probabilities, evaluate the appropriateness of the chosen distribution, make survival, cumulative failure, hazard and probability plots, etc. In situations when no distribution will fit your data, Minitab provides nonparametric estimates of survival probabilities, hazard function, etc.

The failure times can be interval-censored data or warranty data.

Interval-censored data

Your data may be right, left, or interval censored, in addition to exact failure times. You may have multiple failure modes. You can have up to 50 input columns. For multiple samples in the same columns, you can have up to 50 samples.

The first K/2 columns represent start times and the remaining columns represent end times, where K is the number of columns on the command line. Therefore, `LTABLE` will perform K/2 analyses. Each analysis provides identical results to the results you would get if you ran each set of variables separately.

Observation	Variables	Meaning
Exact failure time	Start = End = TIME	Exact failure time = TIME
Right censored	Start = TIME, End = *	Right censored time = TIME
Left censored	Start = *, End = TIME	Left censored time = TIME
Interval censored	Start = TIME1, end = TIME2	Interval censored = (TIME1, TIME2)

Warranty data

To use warranty data, specify the `WDATA` subcommand after the main command. The data have two valid arrangements. In both arrangements, warranty returns are in a triangular matrix. The quantities of units that ship are in either their own column or in the first row of the columns that contain the warranty returns. Each arrangement allows only 1 sample at a time. Use the `STWDATA` subcommand to store the warranty data as interval-censored data to use with the `WPREDICTIONS` command.

Shipment quantities are in a column

Enter the column of shipment quantities first after `LTABLE`. Enter the remaining columns in chronological order from earliest to latest. For example, for the following data, enter

```
LTABLE C2-C6;
WDATA;
```

In this worksheet, the number of items shipped in each year is recorded in a single column, with each row representing a different year. The number of items returned each year is recorded in a separate column. For example, in 2011, 15,000 items were shipped. 78 of those items were returned in 2012, 130 of those items were returned in 2013, and 200 of those items were returned in 2014.

C1	C2	C3	C4	C5	C6
Year Shipped	Quantity Shipped	Returned 2011	Returned 2012	Returned 2013	Returned 2014
2010	10000	50	60	100	300
2011	15000		78	130	200
2012	23000			275	600
2013	32000				400

Shipment quantities are in the first row of the return columns

Enter all the columns in chronological order. For example, for the following data, enter

```
LTABLE C1-C4;
WDATA;
SHPROW;
```

In this worksheet, the first row indicates the total number of items shipped in each year. Each column represents a different year of shipment. Each cell of a column indicates the number of items returned in each subsequent time period. For example, 23,000 items were shipped in 2012. In the following year, 275 of those items were returned. In the year after that, 600 were returned.

C1	C2	C3	C4
Shipped 2010	Shipped 2011	Shipped 2012	Shipped 2013
10000	15000	23000	32000
50	78	275	400
60	130	600	
100	200		

C1	C2	C3	C4
Shipped 2010	Shipped 2011	Shipped 2012	Shipped 2013
	300		

WDATA

Specifies that the columns arguments for the main command contain warranty data.

SHPROW

Indicates that you entered shipment values in the first row of each failure/return column. Without this subcommand, Minitab assumes the shipment values reside in a separate column.

NOPARAMETRIC

When no distribution fits your data, Minitab provides nonparametric estimates of survival probabilities, hazard function, and so on.

Nonparametric estimation methods are Kaplan-Meier and Actuarial methods.

FREQUENCY C...C

Specifies the number of occurrences of a row of data. Frequency columns are useful for data where you have large numbers of observations with common failure and censoring times. Use this subcommand only with interval-censored data.

BY C

For stacked, interval-censored data, C is the variable that defines the groups. Use one By variable to represent your groups. Use this subcommand only with interval-censored data.

Distributions

Use these subcommands to analyze data assuming one of these distributions.

SEXTREMEVALUE

Smallest extreme value distribution.

WEIBULL

Weibull distribution.

TWEIBULL

Weibull distribution.

EXPONENTIAL

Exponential distribution.

TEXPONENTIAL

2[parameter exponential distribution.

NORMAL

Normal distribution.

LNORMAL

Lognormal (base e) distribution.

TLNORMAL
Lognormal (base e) distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

Failure modes

Failure modes are compatible with interval-censored data.

FMODE C or C...C
Estimates overall system reliability and component reliability when multiple causes of failure are present. Specify a single column for all samples or a column for every sample. When you specify a column for every sample, repeat columns are valid. Specify the columns in the same order as the samples for the main command.

KEEP K...K
Specifies which failure modes to include in the analyses.

There may be times when you are only interested in analyzing some of the possible failure modes. For example, you may wish to analyze the overall system reliability with only 1 of the 3 possible failure modes. Specify levels from the column of failure modes to include only those modes in the analysis.

- If the column is numeric, then specify numbers that represent the failure modes.
- If the column is text, then specify text labels that represent failure modes.

ELIMINATE K...K
Specifies which failure modes to exclude from the analyses.

There may be times when you are only interested in analyzing some of the possible failure modes. For example, you may wish to analyze the overall system reliability with only 2 of the 3 possible failure modes. Specify levels from the column of failure modes to exclude those modes from the analysis

- If the column is numeric, then specify numbers that represent the failure modes.
- If the column is text, then specify text labels that represent failure modes.

MIDPOINT
When you calculate right-censored observations for other failure modes when data are interval censored, use MIDPOINT for the midpoint of the interval.

RPOINT
When you calculate right-censored observations for other failure modes when data are interval censored, use RPOINT for the right endpoint of the interval.

CDISTRIBUTION K *"text"*...K *"text"*
Specifies which failure modes to change the distribution of, and which distribution to use. For example, to change the distribution of failure mode A to exponential, type "A" "EXPONENTIAL".

Change the distribution of a failure mode from the assumed overall distribution when a particular failure mode is well modeled by a specific distribution. For example, you specify LTEST; WEIBULL, for the overall system, but change the distribution for failure mode 1 to be lognormal and for failure mode 2 to be normal.

Choose from the distributions described above.

Use K to specify the level of a failure mode that you want to change the distribution for, then type the distribution in double quotes. For example `CDISTRIBUTION "Obstruction" "EXPONENTIAL"` specifies to use the exponential distribution for the failure mode with the label *Obstruction*.

RINDANALYSES
Displays results for individual failure modes.

SINDANALYSES
Stores results for individual failure modes.

Parametric estimation methods

Choose either the LSXY or MLE parametric estimation method.

LSXY
Use LSXY to estimate parameters using least squares (failure time (X) on rank (Y)) method.

MLE
Use MLE to estimate parameters using maximum likelihood method.

COMSHAPE
Use COMSHAPE to assume common shape parameter.

COMSCALE
Use COMSCALE to assume common scale parameter.

TIMES K...K

TIMES C
Calculate the survival probabilities or cumulative failure probabilities for times K...K or C.

ECFPROBABILITY
Choose to estimate cumulative failure probabilities instead of survival probabilities.

STIME
Stores the times associated with the survival probabilities or cumulative failure probabilities.

SURVPROBABILITY C...C
Stores the survival probability for **T**. Where **T** is given on the TIME subcommand.

CLSURVPROBABILITY C...C
Stores the upper and lower confidence limits of the survival probability.

CFPROBABILITY C...C
Stores the cumulative failure probability for **T**. Where **T** is given on the TIME subcommand.

CLCFPROBABILITY C...C
Stores the upper and lower confidence limits of the cumulative failure probability.

PTILES K...K

PTILES C
The percents K which define the additional percentiles to be estimated.

You can enter individual percents (0 < P < 100) or a column of percents.

Bayes analysis

Choose to set the shape (SETSHAPE) or scale (SETSCALE) parameters depending on your distribution.

- For the Weibull distribution, estimate the scale parameter while keeping the shape parameter fixed. The number of arguments must be 1 or equal to the number of variables. If there is 1 argument then all distributions will have the same shape parameter.
- For all distributions except the Weibull distribution, estimate the location parameter while keeping the scale parameter fixed. The number of arguments must be 1 or equal to the number of variables. If there is 1 argument then all distributions will have the same scale parameter.

SETSHAPE K...K

Sets the value for the shape parameter of the Weibull distribution. Enter a list of values, K...K, for the shape that is equal in length to the number of samples.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

If the analysis mixes the Weibull distribution with location-scale distributions, then specify all the shape and scale parameters with the same subcommand. Use the order from the failure modes and samples the same as if all the distributions were the same.

SETSCALE K...K

Sets the value for the scale parameter. Enter a list of values, K...K, for the scale that is equal in length to the number of samples.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

If the analysis mixes the Weibull distribution with location-scale distributions, then specify all the shape and scale parameters with the same subcommand. Use the order from the failure modes and samples the same as if all the distributions were the same.

SETTHRESHOLD K...K

Sets the value for the threshold parameter. Enter a list of values, K...K, for the threshold that is equal in length to the number of samples. To set threshold parameters, all the distributions must have a threshold parameter.

For all distributions with threshold parameters (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic) you can fix the threshold parameter.

The threshold parameter you set must be less than the minimum data value.

If the analysis includes failure modes, then specify the parameters for each failure distribution. The order of the failure modes is alphabetical or from the least number to the greatest.

If the analysis includes failure modes and multiple samples, then specify parameters for each combination. First, specify parameters for all the samples for a particular failure mode, then, proceed to the next failure mode. For example, if the data have 2 failure modes are *Obstruction* and *Break* that are the same in 2 samples, then set the parameters in the following order.

Order	Failure mode	Sample
1	*Break*	First
2	*Break*	Second
3	*Obstruction*	First
4	*Obstruction*	Second

Nonparametric estimation methods

Use TURNBULL or ACTUARIAL as the nonparametric estimation method.

TURNBULL

Use the Turnbull method for nonparametric estimates.

ACTUARIAL

Use the actuarial method for nonparametric estimates.

ECFPROBABILITY

Choose to estimate cumulative failure probabilities instead of survival probabilities.

STIME C

Stores the times associated with the nonparametric survival probabilities or cumulative failure probabilities.

SURVPROBABILITY C...C

Stores the nonparametric survival probability for **T**. Where **T** is given on the TIME subcommand.

SESURVPROBABILITY C...C

Stores the standard error of the nonparametric survival probability.

CLSURVPROBABILITY C...C

Stores the upper and lower confidence limits of the nonparametric survival probability.

CFPROBABILITY C...C

Stores the nonparametric cumulative failure probability for **T**. Where **T** is given on the TIME subcommand.

SECFPROBABILITY C...C

Stores the standard error of the nonparametric cumulative failure probability.

CLCFPROBABILITY C...C
Stores the upper and lower confidence limits of the nonparametric cumulative failure probability.

HAZARD C...C
Stores hazard estimates. These values are available with the acturial estimation method.

HTIME C...C
Stores times associated with hazard estimates. These values are available with the acturial estimation method.

Confidence intervals

Confidence intervals are incompatible with least squares estimation for a parametric analysis.

CONFIDENCE [K]
Specifies the percent confidence, K, for your confidence intervals.

LOWCI
Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI
Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI
Calculates two-sided confidence intervals for parameter estimates and percentiles.

Parameter tests

TVSS K
Tests whether the shape (Weibull) or scale (other distributions) is equal to K (K > 0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the shape parameter equals 2 for two different failure modes, specify `TVSS 2 2;`

TVSL K
Tests whether the scale (Weibull or exponential) or location (other distributions) is equal to K (K > 0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the scale parameter equals 2 for two different failure modes, specify `TVSL 2 2;`

TVTHRESHOLD K
Tests whether the threshold (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic) is equal to K (K>0).

For an analysis with multiple failure modes, specify K for each failure mode. For example, to test that the threshold parameter equals 2 for two different failure modes, specify `TVTHRESHOLD 2 2;`

TESS
Tests for equal shape (Weibull) or scale (other distributions).

TESL
Tests for equal scale (Weibull or exponential) or location (other distributions).

TETHRESHOLD
Tests for equal threshold (three-parameter Weibull, two-parameter exponential, three-parameter lognormal, and three-parameter loglogistic).

Graph options

CI

Displays confidence intervals on probability and survival plots.

PPLOT

Displays a probability plot.

DCPPLOT

Specifies to include symbols for right-censored and left-censored observations on the probability plot.

SPLOT

Displays a parametric or nonparametric survival plot. The nonparametric survival plot is a step function, while the parametric survival plot is a smooth curve.

CFPLOT

Displays a parametric or nonparametric cumulative failure plot. The nonparametric cumulative failure plot is a step function, while the parametric cumulative failure plot is a smooth curve.

HPLOT

Displays a parametric or nonparametric hazard plot. The nonparametric hazard plot is a step function, while the parametric hazard plot is a smooth curve.

PANEL

Use PANEL and SEPARATE to perform hypothesis tests on the distribution parameters.

PANEL displays each graph of different variables or By variable in a separate panel of the graph.

SEPARATE

Use PANEL and SEPARATE to perform hypothesis tests on the distribution parameters.

SEPARATE displays each graph of different variables or By variable in a separate graph.

XMINIMUM K

For parametric distribution analysis, specify the minimum and maximum X scale for the probability, survival and hazard plots. For nonparametric distribution analysis, specify the minimum and maximum X scale for the survival and hazard plot.

Use XMINIMUM to specify the X-axis scale minimum.

XMAXIMUM K

For parametric distribution analysis, specify the minimum and maximum X scale for the probability, survival and hazard plots. For nonparametric distribution analysis, specify the minimum and maximum X scale for the survival and hazard plot.

Use XMAXIMUM to specify the X-axis scale maximum.

XLABEL K

XLABEL *"text"*

Use XLABEL to specify an x-axis label for the graph. When you omit this subcommand, Minitab displays a default label.

Results for parametric and nonparametric analyses

NODEFAULTS

Specifies that no default tables and graphs will be displayed.

TMETHOD

Displays the method table. The table includes information about the estimation method, the number of missing rows, frequency, and any failure modes.

TCENSOR

Displays the censoring table. The table provides information about the censoring scheme, such as any censoring value and the numbers of rows with each type of censoring.

TCDISTRIBUTION

Displays the characteristics of the variable table. For a parametric analysis, the table includes point estimates, standard errors, and confidence intervals for the characteristics of the sample distribution. The characteristics include the Mean Time To Failure (MTTF), standard deviation, median, first quartile, third quartile, and interquartile range. For the least squares (failure time(X) on rank(Y)) estimation method the table provides only point estimates.

For a nonparametric analysis, the table displays point estimates, standard errors, and confidence intervals for the median time to failure. The table is available for the actuarial estimation method.

TPROBABILITY

Displays a table of survival or cumulative failure probabilities. The table includes estimates and confidence intervals of survival or cumulative failure probabilities. For the least squares (failure time(X) on rank(Y)) estimation method the table provides only point estimates of probabilities.

TWDINFO

Displays the columns where analysis stores data transformed from warranty data to interval-censored data. The table is available only when the analysis uses the `STWDATA` subcommand to store the transformed data.

Results for a parametric analysis

STEP

Displays the iteration history table. The table includes the log-likelihood at each iterative step of the optimization algorithm that estimates the distribution parameters.

TPARAMETER

Displays the parameter estimates table. The table includes point estimates for distribution parameters. For the maximum likelihood estimation method, the table also includes standard errors and confidence intervals.

TGOODNESS

Displays the goodness-of-fit table. For maximum likelihood estimation, the table includes the adjusted Anderson-Darling statistic. For least squares estimation the table includes the adjusted Anderson-Darling statistic and the correlation coefficient.

TPERCENTILE

Displays the table of percentiles. The table incldues the estimated percentiles of the distribution. For the maximum likelihood estimation method, the table also includes standard errors and confidence intervals of the percentiles.

TTESTS

Displays the results of the tests of the parameters. The table is available when the analysis includes the subcommand for tests.

Results for a nonparametric analysis

TTURNBULL

Displays a table of Turnbull estimates. The table is avaialble with the Turnbull estimation method. The table includes estimated failure probabilities and standards errors for the intervals in the data.

TATIMES

Displays a table of additional time from time T until 50% of running units fail. The table is available with the actuarial estimation method. The table includes point-estimates of additional survival time from the interval start times. The table also includes standard errors and confidence intervals.

TACTUARIAL

Displays a table of actuarial estimates. The table is available with the acturial estimation method. The table includes estimated conditional failure probabilities and standards errors associated with the intervals in the data.

THAZARD

Displays the hazards and densities table. The table is available with the acturial estimation method. The table includes values of the nonparametric hazard function and the nonparametric density function.

Estimate parameters

START C...C

Specifies the initial values for the estimated parameters. If one column is given then the starting estimates apply for all variables. Otherwise, a column must correspond to each variable.

ITERATION K

Changes the maximum number of iterations that Minitab will perform to reach convergence. The default value is 20. LTEST obtains maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

Storage of fitted distributions

PPERCENTS C

Stores the percents associated with the percentiles.

PERCENTILE C...C

Stores the percentiles.

SEPERCENTILE C...C

Stores the standard error of the percentiles.

CLPERCENTILE C...C

Stores the upper and lower confidence limits of the percentiles.

Storage of estimates

PARAM C...C

Stores the estimated parameters for the distribution down the column.

SSEPARAM C...C

Stores the standard errors of the estimated parameters for the distribution down a column.

CLPARAM C...C
Stores the confidence limits for the estimated parameters for the distribution down a column.

VCPARAM *M...M*
Stores the variance covariance matrix for the estimated parameters.

LOGLIKELIHOOD K...K
Stores the last log-likelihood.

Storage of data

STWDATA
Stores the transformed warranty data in 3 columns after the last column in the worksheet. The first column stores the start times, the second column stores the end times, and the last column stores the frequencies. These columns, containing interval censored data, are the input columns for the Warranty Prediction command.

Subcommand

CTOLERANCE K K
Changes the convergence criteria. Both Ks must be positive numbers.

For the first K, convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than **K**. The default is 0.0001.

For the second K, convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than **K**. The default is 0.0001.

Growth Curves

PGROWTH: Session command for performing a parametric analysis using a parametric growth curve

PGROWTH C...C
Fits a power-law process to each set of recurrence or failure times in C.

Use parametric growth curve (PGROWTH) to perform a parametric analysis on a repairable system. Use either a power-law process or a homogeneous Poisson process to estimate a growth curve of the mean cost of maintaining a system or the mean number of repairs over time.

SEND
Specifies that the response variables are interval failure or retirement times. If you use the SEND subcommand, you must have two response columns for each group. For each pair of response columns, the requirements are as follows:

- The first response variable listed is the start time.
- The second response variable listed is the end time.

BY C

For stacked data, C is the variable that defines the groups. In this case, only one variable (or a pair of variables with the SEND subcommand) is allowed on the PGROWTH command line.

System information

SYSTEM C...C

Identifies a system within a sample when you have multiple systems.

FREQ C...C

Specifies the number of failures recorded at a failure time or interval.

Retirement

RETIREMENT C...C

Put in column C an indicator variable indicating which systems are retired. By default, the lowest numeric value or the lowest ASCII value (text columns) is the retirement value. Retirement columns must contain no more than two distinct values.

RVALUE K

You can specify which value is the retirement value by using the RVALUE subcommand.

TIMETRUNCATE

Specifies that the largest time for each system is the retirement time.

Estimate

By default, Minitab uses maximum likelihood estimation. Use the following subcommands to select a different estimation method.

CMLE

Conditional maximum likelihood estimation. You cannot use CMLE with the SEND or SETSHAPE subcommands.

LSXY

Least squares estimation.

Confidence intervals

CONFIDENCE [K]

Specifies the percent confidence, K, for your confidence intervals.

LOWCI

Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI

Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI

Calculates two-sided confidence intervals for parameter estimates and percentiles.

Graphs

GEVENTPLOT
Plots events for all systems.

DFREQ
Displays the frequency information on the event plot.

GDUANEPLOT
Creates a Duane plot.

GTTTPLOT
Plots the total time on test.

GMCFPLOT
Plots the mean cumulative function.

GONE
By default, Minitab displays plots averaged over all systems. GONE displays one plot per system instead. This subcommand does not apply to the event plot.

PANEL
Displays graphs of different variables or by levels on the same graph in separate panels. By default, multiple plots are overlaid on the same graph.

SEPARATE
Displays graphs of different variables or by levels in separate graphs. By default, multiple plots are overlaid on the same graph.

CI
Includes confidence intervals on the mean cumulative function plot, Duane plot, and TTT plot.

XMINIMUM K
Specifies the X-axis scale minimum.

XMAXIMUM K
Specifies the X-axis scale maximum.

XLABEL K
Labels the X-axis.

Results

STEP
Displays the estimated shape, scale, and log-likelihood at each iteration of the parameter estimation process.

Options

SETSHAPE K
Specifies a known value for the shape parameter instead of estimating it from the data.

START K...K
Specifies the initial value for the shape parameter.

ITERATION K
Specifies the maximum number of iterations to use. The default number of iterations is 20.

Storage

MCF C...C
Stores the mean cumulative function.

CLMCF C...C
Stores the confidence limits for the mean cumulative function.

ROCOF C...C
Stores the rate of occurrence of failures.

CLROCOF C...C
Stores the confidence limits for the rate of occurrence of failures.

PARAMETERS C...C
Stores the estimates of the shape and scale

SSEPARAMETERS C...C
Stores the standard errors of the shape and scale.

CLPARAMETERS C...C
Stores the confidence limits for the shape and scale.

NGROWTH: Session command for performing a nonparametric analysis using a nonparametric growth curve

NGROWTH C...C
Creates a nonparametric growth curve for repair times in C.

Use nonparametric growth curve to estimate a growth curve of the mean cost of maintaining a system or the mean number of repairs over time without assuming a distribution.

SEND
Specifies that the response variables are interval failure or retirement times. If you use the SEND subcommand, you must have two response columns for each group. For each pair of response columns, the requirements are as follows:

- The first response variable listed is the start time.
- The second response variable listed is the end time.

BY C
For stacked data, C is the variable that defines the groups. In this case, only one variable (or a pair of variables with the SEND subcommand) is allowed on the PGROWTH command line.

System information

SYSTEM C...C
Identifies a system within a sample when you have multiple systems.

NSYSTEMS C...C
Specifies the number of systems entering each interval for grouped data.

Retirement

RETIREMENT C...C
Put in column C an indicator variable indicating which systems are retired. By default, the lowest numeric value or the lowest ASCII value (text columns) is the retirement value. Retirement columns must contain no more than two distinct values.

RVALUE K
You can specify which value is the retirement value by using the RVALUE subcommand.

TIMETRUNCATE
Specifies that the largest time for each system is the retirement time.

Cost/Frequency

COST C...C
Specifies a cost or frequency column. You can specify one of the following:

- The total cost for all repairs for an event
- The number of items replaced (frequency)

Graphs

GMCFPLOT
Plots the mean cumulative function.

GDMCPLOT
Plots the mean cumulative difference function. This plot is only available when there are two or more variables (with no SEND subcommand), two or more pairs of variables (with SEND subcommand), or a BY variable.

GEVENTPLOT
Plots events for all systems.

DCOST
Displays the cost variable information on the event plot.

PANEL
Displays graphs of different variables or by levels on the same page in separate panels. By default, multiple plots are overlaid on the same graph page.

SEPARATE
Displays graphs of different variables or by levels on separate graph pages. By default, multiple plots are overlaid on the same graph page.

CI
Includes confidence intervals on the mean cumulative function plot, Duane plot, and TTT plot.

XMINIMUM K
Specifies the X-axis scale minimum.

XMAXIMUM K
Specifies the X-axis scale maximum.

XLABEL K
Labelss the X-axis.

Options

CONFIDENCE [K]
Specifies the percent confidence, K, for your confidence intervals.

LOWCI
Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI
Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI
Calculates two-sided confidence intervals for parameter estimates and percentiles.

Storage

Stores estimates related to the mean cumulative function and the mean cumulative difference function.

MCF C...C
Stores the mean cumulative function.

TMCF C...C
Stores the times for the mean cumulative function.

SEMCF C...C
Stores the standard error of the mean cumulative function.

CLMCF C...C
Stores the confidence limits for the mean cumulative function.

DMCF C...C
Stores the difference in the mean cumulative function.

TDMCF C...C
Stores the times for the difference in the mean cumulative function.

SEDMCF C...C
Stores the standard error of difference in mean cumulative function.

CLDMCF C...C
Stores the confidence limits for difference in mean cumulative function.

Regression with Life Data

LREGRESSION: Session command for performing a regression analysis when the error distribution is Weibull, smallest extreme value, exponential, log-normal, normal, logistic, or log-logistic

LREGRESSION C = C...C

Use LREGRESSION to perform a regression analysis when the error distribution is Weibull, smallest extreme value, exponential, log-normal, normal, logistic, or log-logistic. The data may be right, left or interval censored in addition to exact failure times. The regression equation can have terms which are either continuous or discrete. The command finds maximum likelihood estimates for the parameters using a modified Newton-Raphson technique. Also, you can estimate percentiles, survival probabilities, cumulative failure probabilities, evaluate the appropriateness of the chosen distribution, etc.

The left hand side of the equation specifies the response variables. If SEND is used then the first half of the columns are the START variables while the second half of the columns are the END variables. If the SEND subcommand is not used then each column defines a separate problem. LREGRESSION will do a separate analyses for each set of response variables. Each analyses provide identical results as if you would have run each of the problems separately.

FREQUENCY C...C

Specifies the number of occurrences of a row of data.

FACTORS C...C

Specifies a factor in the model.

Relationships

Specifies the relationship between failure times or log failure times and the accelerating variable.

ARRHENIUS

Arrhenius relationship.

INVTEMP

Inverse temperature relationship.

LN

Power or natural log relationship.

Distributions

Specifies the distribution failure time or log failure time.

WEIBULL

Weibull distribution.

EXPONENTIAL

Exponential distribution.

SEXTREMEVALUE

Smallest extreme value distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

NORMAL
Normal distribution.

LNORMAL
Log-normal (base e) distribution.

Censoring

CENSOR C...C
Indicates which observations are censored. The censored observations are assumed to be right censored while the uncensored observations are considered exact failure times.

CVALUE K
By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value. You can specify which value is the censoring value by using the CVALUE subcommand.

SEND
Specifies that the response variables are in an arbitrary censoring format. You need to enter the response variable into a start and end column.

Estimation

ESTIMATE [K...K]

ESTIMATE [C...C]
Estimates percentiles, survival probabilities, or cumulative failure probabilities for new predictors.

TIMES
Estimates survival or cumulative failure probabilities.

PTILES
The percents K which define the percentiles to be estimated for each new observation.

PERCENTILE
Stores the percentiles.

SEPERCENTILE
Stores the standard error of the percentiles.

CLPERCENTILE
Stores the lower and upper confidence limits for the percentiles.

ECFPROBABILITY
Choose to estimate cumulative failure probabilities instead of survival probabilities.

SURVPROBABILITY
Stores the survival probabilities.

CLSURVPROBABILITY
Stores the lower and upper confidence limits for the survival probabilities.

CFPROBABILITY
Stores the cumulative failure probabilities.

CLCFPROBABILITY
Stores the lower and upper confidence limits for the cumulative failure probabilities.

CONFIDENCE [K]
Specifies the percent confidence, K, for your confidence intervals.

LOWCI
Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI
Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI
Calculates two-sided confidence intervals for parameter estimates and percentiles.

Graphs

RPLOT
Displays a relation plot.

PTILES K...K

PTILES C
Displays the percents associated with the percentiles to be plotted on the relation plot.

PFAILURES
Displays failure times on the relation plot. For exact failure times, the value is plotted while for an interval censored observations the average of the start and end value is plotted.

DVALUE K
Displays the design value K on the relation plot.

ALLCI
Adds confidence intervals on all percentiles on the relation plot.

ONECI
Adds confidence intervals on the middle percentile on the relation plot.

AFPPLOT
Displays a probability plot for each level of the accelerating variable based on the fitted model.

ONECI
Adds a confidence interval for the design value on the probability plot for each level of the accelerating variable based on the fitted model.

ALLCI
Adds a confidence interval for all levels on the probability plot for each level of the accelerating variable based on the fitted model.

AIPPLOT
Displays a diagnostic probability plot for each accelerating variable based on individual fits.

SPPLOT
Displays a diagnostic probability plot for the standardized residuals.

EPPLOT
Displays a diagnostic exponential probability plot for the Cox-Snell residuals.

CI
Display confidence intervals on diagnostic plots.

Results

BRIEF K
Controls the amount of output.

Value of K	What Minitab displays
0	Displays no output, but creates all graphs and stores all specified values.
1	Displays response information, censoring information, regression table, and log-likelihood.
2	Displays table of percentiles and/or cumulative probabilities, in addition to the BRIEF 1 output.
3	Displays list of factor level values, tests for terms with more than 1 degree of freedom, in addition to the output from BRIEF 1 and 2.

STEP
Displays the log-likelihood at each iteration of the parameter estimation process.

Options

ITERATION K
Changes the maximum number of iterations that Minitab will perform to reach convergence. The default value is 20. Minitab's life regression command obtains maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

START
Specifies the initial values for model parameters. The column containing the initial values must have one row for each estimated coefficient in the model. For LREGRESSION, the value in the first row is assigned to the constant, the value in the second row is assigned to the first predictor entered on the main command line, the value in the third row is assigned to the second predictor, etc. Finally, the last value is assigned to the scale parameter if one exists.

You can enter one column of values or a number of columns equal to the number of response variables. There must be one row for each degree of freedom used.

SETS K

SETS K...K
For all distributions except the exponential and Weibull distributions, you can fix the scale parameter and estimate the regression coefficients. The number of arguments must be 1 or equal to the number of response variables. If there is 1 argument then all distributions will have the same scale parameter.

For the Weibull distribution, you can fix the shape parameter and estimate the regression coefficients. The number of arguments must be 1 or equal to the number of response variables. If there is 1 argument, then all distributions will have the same shape parameter.

CTOLERANCE K K

Changes the convergence criteria. Both arguments must be positive numbers.

- Convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than the first K. The default is .0001.
- Convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than the second K. The default is .0001.

TOLERANCE K

Forces Minitab to keep a predictor in the model which is highly correlated with another predictor. Lowering tolerance by giving very small argument values can prevent Minitab from eliminating problematic predictor columns from the model. K is 1 – R-squared, where R-squared is the value resulting from regressing one predictor on the remaining predictors. The default is 1E-18 and the value must be positive.

REFERENCE C K...C K

Changes the reference for factor columns where C is the factor column and K is the reference level.

Storage

RESIDUAL C...C

Stores the ordinary residuals.

SRESIDUAL C...C

Stores the standardized residuals.

CRESIDUAL C...C

Stores the Cox-Snell residuals.

COEFFICIENT C...C

Stores the coefficients for the fitted model down a column in the order they appear in the model.

SECOEFFICIENT C...C

Stores the standard errors for the coefficients.

CLCOEFFICIENT C...C

Stores the lower and upper confidence limits for the coefficients.

VCCOEFFICIENT *M...M*

Stores the variance-covariance matrix of the estimated coefficients.

LOGLIKELIHOOD K...K

Stores the last value of the log-likelihood.

Cox Regression

COXREGRESSION: Session command for performing a Cox regression analysis

COXREGRESSION

Use COXREGRESSION to describe the relationship between predictors and survival time. The analysis uses both fixed predictors and time-dependent predictors. To use time-dependent predictors, arrange the data in the counting process form. The regression equation can have terms which are either continuous or categorical. You can include interaction and polynomial terms and perform stepwise selection of terms.

Key results of comparative studies that use Cox regression often report relative risks for predictors and display graphs of the survival experience of subjects under different treatments. For example, a study on a cancer treatment concludes that the relative risk for two groups is 4, which means that the patients in one group are cancer-free at 4 times the rate of patients in the other group over the study period. Minitab displays relative risks for each variable so that you can easily assess their effect on the frequency of the event.

Specify the response for a model with fixed predictors only

For a model with fixed predictors only, the analysis requires `RESPONSE`.

RESPONSE C

The observed time for each observation. This value is either the time until the event of interest or the censoring time. The column must be numeric.

ENTRYTIME C

The entry time for each observation. The column must be numeric and the same length as the column you specified with `RESPONSE`. The entry times must be less than or equal to the response times. If you do not specify a column, Minitab uses 0 for the entry time for all observations.

You usually use entry times to model left-truncated data in models with fixed predictors. Left truncation occurs when potential subjects of a study are observed at a time after the start of the study.

CENSOR C

Indicates which observations are censored. For a censored observation, the value in the response column is the latest time for the row even though the event did not occur.

CVALUE K

Specifies which value indicates a censored observation. By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value.

Specify the response for a model with time-dependent predictors or recurrent events

For a model with time-dependent predictors or recurrent events, the analysis requires `CPIF`, `RESPONSE`, and `ENTRYTIME`.

CPIF

Indicates that the data are in a counting process form.

RESPONSE C

The end time for each interval of observation. This value is either the time until the event of interest or the censoring time. The column must be numeric.

ENTRYTIME C

The start time for each interval of observation. The column must be numeric and the same length as the column you specified with RESPONSE. The start time must be less than or equal to the end time for that row. If you do not specify a column for the entry times, Minitab uses 0 for the entry time for all observations.

You usually use entry times to model left-truncated data in models with fixed predictors. A time is left-truncated when the subject enters after time 0.

CENSOR C

Indicates which observations are censored. The end of the time interval for a censored observation indicates that the event of interest did not occur yet. Because the counting process form allows recurrence data, censored intervals for a subject can appear after an uncensored interval.

CVALUE K

Specifies which value indicates a censored observation. By default, the lowest numeric value or the lowest ASCII value for text columns is the censoring value.

Model

CONTINUOUS C...C

Specifies the continuous predictors if you have any. The column(s) must be numeric or date/time and must match the length of the response column.

CATEGORICAL C...C

Specifies the categorical predictors if you have any. The column(s) can be numeric, text, or date/time and must match the length of the response column.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. You can only repeat continuous predictors in a cross-term. Nested terms are not allowed. The model can be nonhierarchical.

STRATIFY C [C]

Fit a stratified model. Each unique combination of values in the columns that you specify defines a stratum. A stratified model estimates a baseline hazard rate for each strata but uses the same estimates for the effects of the predictors. You can have up to two stratification variables.

Coding categorical predictors

REFERENCE C K ... C K

Changes the default coding for the categorical predictor columns. To change the default reference factor level, specify the factor column followed by the reference level. (You must enclose text and date/time levels in double quotation marks.) You can assign a reference level only when you use the binary coding (1, 0) scheme.

EFFECT

Specifies the effect coding (−1, 0, +1) scheme for categorical predictors.

Standardizing the continuous predictor

Use this set of subcommands to standardize the continuous predictors in your model. If you specify both SCALE and LOCATION Minitab performs the subtraction before the division. LEVELS is mutually exclusive with LOCATION and SCALE.

If you do not specify LOCATION, SCALE, or LEVELS Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Coding of Predictors**.

LOCATION [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by subtracting a constant from each predictor. If you do not specify any arguments, the mean of each predictor column is subtracted. K specifies to subtract a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

SCALE [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by dividing each predictor by a constant. If you do not specify any arguments, each predictor column is divided by the standard deviation. K specifies to divide by a constant. If you specify arguments, the number of arguments must match the number of continuous predictors.

LEVELS [K...K]

Specifies that the analysis is to be performed on coded continuous predictors by DOE-type coding for the specified low and high levels K K&K K. The number of arguments must be twice the number of continuous predictors.

Options

Additional options that you can specify for the analysis.

BRESLOW

Use the Breslow method for handling ties. If you do not specify this subcommand, Minitab uses the Efron method.

CLUSTERID[C]

Perform the analysis using the WLW robust covariance matrix for the parameter estimates. If you specify a column, Minitab uses it to identify groups of correlated observations due to the study design. For example, in recurrent event models where each subject can experience the event multiple times the observations within the same subjects are correlated. If you specify a column, Minitab calculates the robust covariance to account for the presence of clustered observations. Rows with the same value are clustered observations. If you do not specify a column, the effect is the same as if you use a column with a different value in every row.

All the tests and confidence intervals in the analysis are based on the robust covariance matrix if you specify CLUSTERID. The input column can be numeric, text or date/time. Missing row values do not cause the rows to be missing in the analysis. The missing rows are grouped together when calculating the robust variance-covariance matrix.

WALD

Use the Wald test for the ANOVA table. You can specify only one of WALD, LRT, and SCOTEST. If you do not specify the type of test, then the analysis uses the Wald test for the ANOVA table. If you specify CLUSTERID, then the local Wald tests in the ANOVA table use the robust variance-covariance matrix instead of the model-based variance-covariance matrix.

LRT

Use the likelihood ratio test for the ANOVA table. If you don't specify this subcommand, Minitab uses the Wald test. If you specify CLUSTERID, then the ANOVA table displays Wald tests because likelihood ratio tests assume that the observations within clusters are independent.

SCOTEST

Use the score test for the ANOVA table. If you don't specify this subcommand, Minitab uses the Wald test. If there are no tied event times, then the score test is identical to the log-rank test. If you specify CLUSTERID, then the ANOVA table displays Wald tests because likelihood ratio tests assume that the observations within clusters are independent.

CASEID C

Specifies a column to identify the subjects in the study. Use this subcommand only if you also specify CPIF. The column can be numeric, text, or date/time. You can identify a subject using a case number, ID, or name. Minitab groups missing rows together to calculate the per case diagnostics.

In a counting process input format, a single subject can have multiple rows of diagnostic statistics. If you specify a case-identification column, Minitab provides a single diagnostic statistic per case by summing the statistics over the multiple rows pertaining to the same case.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter CONFIDENCE 90. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Stepwise

FINFORMATION, STEPWISE, FORWARD, and BACKWARD perform a stepwise regression procedure to fit the model. No arguments are needed for these subcommands. If you specify CLUSTERID, only BACKWARD uses the robust variance-covariance matrix at each step. FORWARD always use the model-based variance-covariance matrix, even if you specify CLUSTERID.

FINFORMATION

Specifies a stepwise model selection procedure that uses forward information criteria selection. Use AICCORRECTED or BICRITERION to specify which information criterion to use to select the final model. If you do not specify a criterion, Minitab uses AICCORRECTED.

The forward information criteria procedure adds the term with the lowest p-value to the model at each step. If you do not include subcommands about hierarchy, FINFORMATION adds 1 term at a step and maintains model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

Minitab calculates the information criteria for each step.

In most cases, the procedure continues until one of the following conditions occurs:

- The procedure does not find an improvement of the criterion for 8 consecutive steps.
- The procedure fits the full model.
- The procedure fits a model that leaves 1 degree of freedom for error.

If you specify settings for the procedure that require a hierarchical model at each step and allow only one term to enter at a time, then the procedure continues until it either fits the full model or fits a model that leaves 1 degree of freedom for error. Minitab displays the results of the analysis for the model with the minimum value of the selected information criterion, either the corrected Akaike's Information Criterion (AICc) or the Bayesian Information Criterion (BIC).

AICCORRECTED

Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the final model.

BICRITERION

Specifies the use of the Bayesian Information Criterion (BIC) to select the final model.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

FORWARD

Specifies a stepwise model selection procedure that uses forward selection. If you do not include subcommands about hierarchy, STEPWISE and FORWARD add 1 term at a step and maintain model hierarchy, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS; SINGLE.
```

BACKWARD

Specifies a stepwise model selection procedure that uses backward elimination. Removes a single term at each step and maintains a hierarchical model, the equivalent of the following:

```
HIERARCHICAL; ALLTERMS; ALWAYS.
```

AENTER K

Specifies the alpha level at which a term is entered into the model. The default is 0.15 for STEPWISE and 0.25 for FORWARD.

AREMOVE K

Specifies the alpha level at which a term is removed from the model. The default is 0.15 for STEPWISE and 0.10 for BACKWARD. For STEPWISE, K must be greater than or equal to K for AENTER.

ENTER *termlist*

Specifies the terms that are contained in the starting model for STEPWISE. The ENTER termlist must be a subset of the TERMS termlist or in the default term list in the design.

FORCE *termlist*

Specifies the terms to be forced in the model. The FORCE termlist must be a subset of the TERMS termlist or in the default term list in the design.

NOHIERARCHICAL

Specifies that the model selection procedure does not consider hierarchy.

HIERARCHICAL

Maintains a hierarchical model in stepwise regression. In a hierarchical model, if a higher-order term is included, all lower-order terms that comprise the higher-order term also appear in the model. For example, a model that includes the interaction term A*B*C is hierarchical if it includes the following main effects and lower-order interactions: A, B, C, A*B, A*C, and B*C.

CATONLY

Specifies that only the categorical terms in the model have to be hierarchical.

ALLTERMS

Specifies that both categorical and continuous terms have to be hierarchical.

ATEND

Specifies that the final step of the stepwise procedure adds terms to make the model hierarchical.

ALWAYS
Specifies that the model is hierarchical at every step.

SINGLE
Specifies that only one term can enter the model at each step. So a higher-order term can enter the model only if the terms that comprise the term are already in the model. For example, the algorithm does not consider the addition of A*B unless A and B are already in the model.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

MULTIPLE
Specifies that multiple terms can enter the model at each step. So a higher order term can enter the model, and the terms that comprise the term enter the model at the same time. For example, if A*B is the most statistically significant term, A*B enters the model. At the same time, A and B enter the model if those terms are not in the model already.

BACKWARDS does not use SINGLE or MULTIPLE because terms only exit the model.

Results

NODEFAULT
Specifies that no default tables or graphs will be displayed

TMETHOD
Displays the method table.

TMSDETAILS
Displays the type of stepwise procedure and the alpha values to enter and/or remove a predictor from the model. Use `FULL` to display the coefficients, p-values, and model summary statistics for each step of the procedure. Use `NOFULL` to hide these statistics. If you do not specify `FULL` or `NOFULL`, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Stepwise**.

TLEVELS
Displays the table of categorical predictors and their levels.

TSTEP
Displays the log-likelihood at each iterative step of the optimization algorithm.

TCENSORING
Displays the table showing the number of uncensored and censored values. If the analysis is stratified, there is a row for each stratum plus a row for totals.

TEQUATION
Displays the regression equation table. Minitab will display up to 50 equations. If you want to see a single equation, rather than a separate equation for each factor level combination, use `SINGLE`. To see the separate equations, use `SEPARATE`. If you do not specify `SINGLE` or `SEPARATE`, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

TANOVA
Displays the ANOVA table.

TSUMMARY
Displays the model summary table.

TGOODNESS

Displays the goodness-of-fit tests table. The table includes the likelihood ratio, Wald, and score testing methods. If you specify CLUSTERID, then the likelihood ratio test results are missing values because the method requires independence of observation within clusters.

TCOEFFICIENTS

Displays the table of coefficients. Use FULL to display the full set of coefficients for categorical predictors. Use NOFULL to show only the linearly independent coefficients. If you do not specify FULL or NOFULL, Minitab uses the preferences set in **File** > **Options** > **Linear Models** > **Display of Results**.

If you specify CLUSTERID, then the statistics are based on the robust variance-covariance matrix.

TRRISK

Displays the relative risks table. Minitab displays a separate table for continuous and categorical predictors.

INCREMENT K...K

Specify the unit of change for the continuous predictors. The number and order of arguments must match the arguments you specify with CONTINUOUS. If you do not specify INCREMENT, the unit of change is 1.

TPHTEST

Displays the tests of proportional hazards table. The table contains tests of linear associations between the scaled Schoenfeld residuals and some function, $g(t)$, of the event times. In practice, the selected function $g(t)$ are the log function, the Kaplan-Meier's survival function, the rank function, or the identity function. You can select these functions with the following subcommands:

LOG

Use the natural log function, $g(t) = \ln(t)$, to transform the event times for the tests of proportional hazards.

KMESTIMATES

Use the Kaplan-Meier's survival function to transform the event times for the tests of proportional hazards. The survival estimates are based on the response time.

RANK

Use the rank function, $g(t) = \text{Rank}(t)$, to transform the event times for the tests of proportional hazards. The rank function is based on the ranks of the event times. Tied values are assigned their average ranks.

IDENTITY

Use the original units for the tests of proportional hazards. Minitab uses this option if you do not specify a function.

TSURVIVAL

Displays the survival function table. Minitab uses the mean of each predictor for continuous predictors and the reference levels for categorical predictors. You cannot specify this subcommand if you also specify CPIF.

LOG

Calculate the confidence intervals based on a log transformation of the survival function. This is the default setting.

CLOGLOG

Calculate the confidence intervals based on the complementary log-log transformation of the survival function.

TEXPAND

Displays the expanded version of tables.

Graphs

GSURVIVAL

Displays the survival plot. By default, Minitab uses the mean of each predictor for continuous predictors and the reference levels for categorical predictors. GSURVIVAL has two optional subcommands to specify new predictor values for the survival plot. You cannot specify this subcommand if you also specify CPIF. The subcommands PANEL, OVERLAY, and SEPARATE only affect the graph if the model is stratified.

CPREDICTOR C...C

Specify a column of predictor values that Minitab uses for the survival function. The number and type of arguments must match the variables in the model. Specify columns with continuous predictors first, then specify columns with categorical predictors. For a categorical predictor, a new value must be one of the levels of the categorical predictor.

KPREDICTORS K...K

Specify individual predictor values that Minitab uses for the survival function. Specify continuous predictors first, then specify categorical predictors. You can repeat this subcommand. If you specify a missing value for a continuous predictor, Minitab uses the mean of that predictors. If you specify a missing value for a categorical predictor, Minitab uses the reference level. If you enter a missing value as a numeric value, enclose it in single quotes. You can enclose text and date/time values in single or double quotes.

PANEL

Displays the survival plots associated with each stratum in a separate panel of the same graph. This is the default setting.

OVERLAY

Overlay the survival plots associated with each stratum on the same graph.

SEPARATE

Displays the survival plots associated with each stratum in a separate graph in the output pane.

GDEVIANCE

Displays a graph of the deviance residuals versus the risk scores. The graph measures the effect of a given subject on the model. You can also use it to detect outliers.

GANDERSEN

Displays the Andersen plot for the stratification variables. Use this graph to check whether the proportional hazards assumption is satisfied for the stratification variables. If the assumption holds, the curves should be straight lines through the origin. If the variables do not need to be in the model, the curves should roughly follow the 45° line. Minitab only displays this plot if the model is stratified.

GARJAS C...C

Display the Arjas plots for the variables that you specify. Use to check for a missing predictor and whether the proportional hazards assumption holds for a predictor. You must specify at least one column. The column can be any data type but must have the same number of rows as the column you specify with RESPONSE. You must categorize a continuous variable before you use it in an Arjas plot.

GVARIABLE C...C

Display graphs of the martingale residuals versus the variables that you specify. Use to assess whether you should add a predictor to the model, or whether you should use a different functional form for an existing predictor. You must specify at least one column. The column must be numeric or date/time data and must have the same number of rows as the column you specify with RESPONSE.

Storage

COEFFICIENT [C]
Stores the coefficients for the fitted model in a column in the order they appear in the model.

SECOEFFICIENT [C]
Stores the standard errors for the coefficients.

SCOVARIANCES *M*
Stores the variance-covariance matrix of the estimated coefficients.

SCORRELATION *M*
Stores the correlation matrix of the estimated coefficients.

STST [C]
Stores the test statistics for the coefficients.

SPVALUES [C]
Stores the p-values for the coefficients.

XMATRIX *M*
Stores the design matrix for the regression model.

MARRESIDUALS [C]
Stores the martingale residuals. If the data is in a counting process format and you specify a case ID variable, then Minitab stores per subject residuals by summing the per observation residuals over the rows pertaining to the same subject. If you do not specify a case ID variable then Minitab stores per observation residuals which are not necessarily per subject residuals.

DEVRESIDUALS [C]
Stores the deviance residuals. If you specify a smaller number of columns, then you get a subset of the parameters. If the data is in a counting process format and you specify a CASEID variable, then Minitab stores per subject deviance residuals. If you do not specify a CASEID variable then Minitab stores per observation deviance residuals which are not necessarily per subject deviance residuals.

SCHRESIDUALS [C...C]
Stores the Schoenfeld residuals. Minitab stores residuals in a column for each parameter in the model. If the data is in a counting process format and you specify a CASEID variable, then Minitab stores per subject residuals by summing the per observation residuals over the rows pertaining to the same subject. If you do not specify a CASEID variable then Minitab stores per observation residuals which are not necessarily per subject residuals.

WSCHRESIDUALS [C...C]
Stores the weighted Schoenfeld residuals. Minitab stores residuals in a column for each parameter in the model. If the data is in a counting process format and you specify a CASEID variable, then Minitab stores per subject residuals by summing the per observation residuals over the rows pertaining to the same subject. If you do not specify a CASEID variable then Minitab stores per observation residuals which are not necessarily per subject residuals.

SCRESIDUALS [C...C]
Stores the score residuals. Minitab stores residuals in a column for each parameter in the model. If the data is in a counting process format and you specify a CASEID variable, then Minitab stores per subject residuals by summing the per observation residuals over the rows pertaining to the same subject. If you do not specify a CASEID variable then Minitab stores per observation residuals which are not necessarily per subject residuals.

DFBETAS [C...C]

Stores the DFBetas, also known as scaled score residuals. Minitab stores residuals in a column for each parameter in the model. If the data is in a counting process format and you specify a CASEID variable, then Minitab stores per subject DFBetas by summing the per observation DFBetas over the rows pertaining to the same subject. If you do not specify a CASEID variable then Minitab stores per observation DFBetas which are not necessarily per subject DFBetas.

LOGLIKELIHOOD [C]

Stores the first and the last value of the log-likelihood.

STETIME [C]

Stores the transformed event times associated with the proportional hazard test. Only specify this sucommand if you also specify TPHTEST.

Algorithm options

TOLERANCE K

Specifies the tolerance level for collinearity and constant check. By default, K = 4 * 2.22e–012.

ITERATIONS K

Specifies the maximum number of iterations. By default, K = 25.

CTOLERANCE K K

Specifies the convergence tolerances used in the Newton-Raphson algorithm for maximizing the log-likelihood. The value you enter for the first K is for the convergence of log-likelihood values. The value you enter for the second K is for the convergence of the estimates. The default for both is 0.0001.

Probit Analysis

PROBIT: Session command for performing probit analysis

PROBIT C [C] = C [C]

Analyzes the relationship between a binary response in C and a single stimulus and optional factor in C C using probit analysis.

Use probit analysis when you want to estimate percentiles, survival probabilities, and cumulative failure probabilities for the lifetime distribution of a system at some stress, and draw probability plots. When you enter a factor and choose a Weibull, lognormal, or loglogistic distribution, you can also compare the potency of the stress under different conditions.

Minitab calculates the model coefficients using a modified Newton-Raphson algorithm. The left-hand side of the equation specifies the response variable. Enter one column if using a response/frequency format and two columns if using success/trial, success/failure or failure/trial format. On the right-hand side of the equation, enter a stress variable and optionally a factor variable.

FREQUENCY C

Specifies multiple occurrences of the same row of data when your data are in response/frequency format.

ST
The format of the response is a column of events (successes) and a column of trials.

SF
The format of the response is a column of events (successes) and a column of failures.

FT
The format of the response is a column of non-events (failures) and a column of trials.

FACTOR C
Specifies a factor in the model.

Distributions

Specifies the tolerance distribution.

WEIBULL
The Weibull distribution.

SEXTREMEVALUE
The smallest extreme value distribution.

LOGISTIC
The logistic distribution.

LLOGISTIC
The loglogistic distribution.

NORMAL
The normal distribution.

LNORMAL
The lognormal distribution.

Estimation

PTILES K...K

PTILES C
Specifies the additional percentiles to be estimated.

You can enter individual percents (0 < P < 100) or a column of percents.

STRESS K...K

STRESS C
Calculates the survival probability for stresses in K or C.

ECFPROBABILITY
Estimates cumulative failure probabilities instead of survival probabilities.

HOSMER K
Specifies the number of groups for the Hosmer-Lemeshow goodness-of-fit test.

NAPPROXIMATION
Specifies a normal approximation confidence interval instead of the default Fiducial confidence intervals.

CONFIDENCE [K]
Specifies the percent confidence, K, for your confidence intervals.

LOWCI
Calculates one-sided, lower confidence intervals for parameter estimates and percentiles.

UPCI
Calculates one-sided, upper confidence intervals for parameter estimates and percentiles.

TWOCI
Calculates two-sided confidence intervals for parameter estimates and percentiles.

Graphs

PPLOT
Displays a probability plot that describes the fitted model.

SPLOT
Displays a survival plot that describes the fitted model.

CFPLOT
Displays a cumulative failure plot that describes the fitted model.

CI
Specifies to include confidence intervals on the probability, survival, and cumulative failure plots.

PPEARSON
Displays a plot of the Pearson residuals versus event probability.

PDEVIANCE
Displays a plot of the deviance residuals versus event probability.

PANEL
Displays graphs of different variables or by levels in separate panels on the same graph.

SEPARATE
Displays graphs of different variables or by levels on separate graphs.

Options

ITERATION K
Changes the maximum number of iterations that Minitab will perform to reach convergence. The default value is 20. Minitab's Probit analysis command obtains maximum likelihood estimates through an iterative process. If the maximum number of iterations is reached before convergence, the command terminates.

START C
Specifies the initial values for model coefficients. The column containing the initial values must have one row for each estimated coefficient in the model. For Probit Analysis, the value in the first row is assigned to the constant, the value in the second row is assigned to the first predictor entered on the main command line, the value in the third row is assigned to the second predictor, and so on.

SETC K
Sets the natural response value to K.

OPTC
Specifies to estimate the natural response from the data. If you also specify SETC, then SETC gives the starting value for the estimation.

REFERENCE C K, ..., C K
Changes the default coding for a response variable and/or factor. Enter the column followed by the reference value.

CTOLERANCE K K
Changes the convergence criteria. Both arguments must be positive numbers.

- Convergence is attained relative to this criterion if from one iteration to the next the change in the log-likelihood is less than the first K. The default is .0001.
- Convergence is attained relative to this criterion if from one iteration to the next the maximum change in the coefficients is less than the second K. The default is .0001.

Results

BRIEF K
Controls the amount of output.

Value of K	What output Minitab displays
0	Minitab displays no output, but displays all requested graphs and performs all requested storage.
1	Displays response information, regression table, test for equal slopes, log-likelihood, multiple DF test and 2 goodness-of-fit tests.
2	Displays estimated distribution parameters, table of percentiles and survival probabilities, in addition to the BRIEF 1 output.
3	Displays characteristic of distribution and Hosmer-Lemeshow goodness-of-fit test, in addition to the output from BRIEF 1 and 2.

STEP
Displays the log-likelihood at each iteration of the parameter estimation process.

Storage of residuals and percentile information

PRESIDUAL C
Stores the Pearson residuals.

DRESIDUAL C
Stores the deviance residuals.

PPERCENT C
Stores the percents associated with the percentiles.

PERCENTILE C
Stores the percentiles.

SEPERCENTILE C
Stores the standard error of the percentiles.

CLPERCENTILE C
Stores the lower and upper confidence limits for the percentiles.

Storage of information for the survival or cumulative failure probabilities

The following subcommands store information related to the survival or cumulative failure probabilities. You can use these subcommands only if the STRESS subcommand is given.

SSTRESS C
Stores the stress values associated with the survival or cumulative failure probabilities.

SURVPROBABILITY C
Stores the survival probabilities.

CLSURVPROBABILITY C
Stores the upper and lower confidence limits of the cumulative failure probabilities.

CFPROBABILITY C
Stores the cumulative failure probabilities.

CLCFPROBABILITY C
Stores the upper and lower confidence limits of the cumulative failure probabilities.

Storage of information for the estimated equation

EPROBABILITY C
Stores the event probabilities.

COEFFICIENT C
Stores the coefficients for the fitted model down a column in the order they appear in the model.

SECOEFFICIENT C
Stores the standard errors of the coefficients down a column in the order that they appear in the model.

NATRESPONSE K
Stores the natural response rate.

SENATRESPONSE K
Stores the standard error of the natural response.

VCCOEFFICIENT *M*
Stores the variance-covariance matrix of the estimated parameters.

LOGLIKELIHOOD K
Stores the log-likelihood from the last iteration.

Warranty Analysis

TWARRANTY: Session command for pre-processing warranty data

TWARRANTY C...C

Pre-processes the warranty data in columns C...C to interval censored data. If you enter shipment values in a column, the first argument corresponds to this shipment column, and subsequent arguments refer to columns of failures or returns, input in chronological order. Alternatively, if you enter shipment values in the first row of each return/failure column, arguments correspond to these columns, input in chronological order.

After pre-processing your warranty data with this command, you can complete your warranty analysis with the Warranty Prediction command. TWARRANTY accepts warranty data in two formats: you can either enter the numbers of items shipped in a single column alongside the triangular matrix of return/failure counts, or you can enter the numbers of items shipped in a row above the triangular matrix containing the return/failure counts.

SHPROW

Indicates that you entered shipment values in the first row of each failure/return column. Without this subcommand, Minitab assumes the shipment values reside in a separate column.

STWDATA C C C

Stores the transformed warranty data in columns C, C, and C. The first column stores the start times, the second column stores the end times, and the last column stores the frequencies. These columns, containing interval censored data, are the input columns for the Warranty Prediction command.

AFTER

Appends the transformed warranty data into three empty columns after the last column used in the current worksheet. This subcommand is not supported in local macros.

WPREDICTIONS: Session command for performing warranty predictions

WPREDICTIONS C C

Performs warranty predictions on your interval censored data in C C. The first argument is the "start time" and the second argument is the "end time".

WPREDICTIONS uses information about current warranty claims to predict the number of claims in the future. Input data must be interval censored; use TWARRANTY command to convert warranty data from the popular triangular matrix format to interval censored format.

FREQUENCY C

Specifies the number of occurrences of a row of the interval censored data. The length of the argument must equal the length of the argument for WPREDICTIONS.

Distributions

Specifies the distribution that the expected failure times are assumed to follow.

SEXTREMEVALUE

Smallest extreme value distribution.

WEIBULL
Weibull distribution.

TWEIBULL
Three-parameter Weibull distribution.

EXPONENTIAL
Exponential distribution.

TEXPONENTIAL
Two-parameter exponential distribution.

NORMAL
Normal distribution.

LNORMAL
Log-normal distribution.

TLNORMAL
Three-parameter log-normal distribution.

LOGISTIC
Logistic distribution.

LLOGISTIC
Log-logistic distribution.

TLLOGISTIC
Three-parameter log-logistic distribution.

Estimation methods

LSXY
Estimates model parameters using the least squares estimation method.

MLE
Estimates models parameters using the maximum likelihood estimation method.

ESTIMATES C
Specifies estimated parameters (historical estimates or estimates obtained from parametric distribution analysis).

Prediction options

PREDICT K
Predicts the number of failures in future time periods 1 to K.

PRODUCTION K...K

PRODUCTION C
Specifies the projected number of new units to be produced during the future time periods for which you want to predict failures. If you enter one constant, it applies to all future time periods. If you enter multiple constants or a column, values will correspond respectively to the future time periods. The number of constants, or the number of values in the column, must equal the argument for the PREDICT subcommand.

WLENGTH K
Specifies the length of the warranty in K units of time. The time units must be the same as the time units of the data itself.

COST K
Specifies the average cost K associated with a failure.

LOWCI
Specifies a one-sided lower confidence bound for the predicted number of failures based on a Poisson approximation.

UPCI
Specifies a one-sided upper confidence bound for the predicted number of failures based on a Poisson approximation.

TWOCI
Specifies a two-sided confidence interval for the predicted number of failures based on a Poisson approximation.

Graph options

GPFPLOT
Plots the predicted failures against their future time periods.

GPCPLOT
Plots the predicted cost of failures against their future time periods.

GCI
Displays Poisson confidence intervals on the predicted failures plot.

Storage options

SPFAILURES C
Stores predicted future failures.

SPCOST C
Stores predicted future costs of failures.

SCLFAILURES C
Stores lower and upper confidence limits for the predicted failures in C and C.

SCLCOST C C
Stores lower and upper confidence limits for the predicted costs in C and C.

Predictive Analytics

CART Regression

RTREE: Session command for creating a regression tree

RTREE

Creates an optimal regression tree and a set of close-to-optimal regression trees. This analysis also provides residual plots, an important variables chart, and other diagnostic measures so that you can assess the validity of your regression model.

The response variable column must be continuous and must contain at least 2 distinct numeric values.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable or the weights column.

RESPONSE C

Enter a column that contains 2 or more distinct values of a continuous response. C must be numeric.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Validation

Use one of the following subcommands to specify the validation method.

- CVTEST
- TEST
- NTEST

If no validation subcommand is issued, then Minitab uses the following defaults, depending on the number of rows:

The default when the number of rows ≤ 5000

CVTEST 10;

RANDOMIZE 12345;

The default when the number of rows > 5000

TEST 0.3;

RANDOMIZE 12345;

CVTEST K

Use K-fold cross-validation as a test method. Specify K number of folds.

K must be an integer between 2 and the sample size; $2 \leq K \leq 1000$.

RANDOMIZE K

Specify the base number for randomly splitting the training data set into K subsets for cross-validation. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to reproduce the default results from the graphical user interace, where 12345 is the default seed. Without the RANDOMIZE subcommand, the analysis uses a randomly selected seed.

CVTEST C

Use K-fold cross-validation as a test method. Enter C to specify the rows for each fold.

C can be numeric or text and must be the same length as the response column.

TEST K

Specify a proportion of rows for the test set. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to reproduce the default results from the graphical user interace, where 12345 is the default seed. Without the RANDOMIZE subcommand, the analysis uses a randomly selected seed.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

NTEST

No model validation.

Use `SSAMPLE C` to store a column that shows the assignments for each row in the validation. When `CVTEST` is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When `TEST` is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the two values.

Options

Use one of the following subcommands to specify the node splitting method.

- `LSE`: Use the Least Squared Errors method to split a node. This is the default method.
  - `MAXRSQ`: Use maximum R^2 from test set to select the optimal tree.
  - `STERROR K`: Use within K standard error of the maximum R^2 to select the optimal tree.

- LAD: Use the Least Absolute Deviation method to split a node.
  - MINMAD: Use minimum relative mean absolute deviation (MAD) from test set to select the optimal tree.
  - STERROR K: Use within K standard error of minimum MAD to select the optimal tree.

If you do not specify the node splitting method, then the analysis uses LSE and STERROR 1.

The following subcommands are available to control tree growth due to sample size or node complexity.

- NOSPLIT K: Do not split a node if the node size is less than K. K=10 is the default.
- NOTERMINAL K: Do not create a terminal node that is smaller than K. K=3 is the default.
- MAXTDEPTH K: Set maximum tree depth to the Kth layer.
- NLIMIT K: Set the maximum number of terminal nodes to K. K=1024 is the default.

SURROGATE K

Specify K surrogates to consider when a specific predictor that splits a node has missing values.

K must be an integer between 0 and the number of predictors – 1, for example:

- K = 0: Specifies no consideration of surrogate variables.
- K = Number of total predictors – 1: Specifies to consider all predictors other than the splitter variable.

If this subcommand is not issued, K=10 is the default value.

WEIGHT C

Specify the case weights in C. The column length must be the same length as the target column length.

Values must be ≥ 0. Assigning a weight of 0 or missing excludes the row from the analysis.

MPENALTY K

Specify a penalty value for a predictor with missing values.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty.
- K = 2: Specifies highest penalty.

.

CPENALTY K

Specify a penalty value for categorical predictors that have many values.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty.
- K = 5: Specifies highest penalty.

Graphs

GTDIAG

Displays the optimal tree diagram.

GRSQ

Displays the R^2 vs number of terminal nodes graph when LSE is issued. If LAD is issued, GRSQ displays relative MAD vs number of terminal nodes graph.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GFIT

Displays fits vs actual response values plot.

RTYPE K

Specifies the type of residual to plot. When K =

- 1: Displays regular residuals.
- 2: Displays percent residuals.

GRBOX

Displays boxplot of the residuals. If the analysis uses a validation method, the graph has separate boxplots for the training data set and the test data set.

SORT

Sorts terminal nodes on the x-axis in mean squared error (MSE) or mean absolute deviation (MAD) ascending order. Otherwise, the terminal nodes are in order by identification number.

GERROR

Displays mean squared error (MSE) or mean absolute deviation (MAD) vs terminal node id plot. Y axis = MSE if LSE is issued; otherwise Y axis = MAD

GBOX

Displays boxplot of response values by terminal node.

GRESIDUAL

Displays residual vs terminal node id plot.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table, which includes the model validation method, the node splitting method, and other information about the analysis.

TRESPONSE

Displays the Response information table, which includes descriptive statistics about the response variable.

TSUMMARY

Displays the Model summary table, which includes the number of predictors, the number of important predictors, the number of nodes, the R^2 statistic, and other information about the model.

TPERROR K

Displays percent of error statistics due to largest residuals. K represents the percent of largest residuals being displayed.

0 < K ≤ 100. K = 20 is the default, which means the table includes statistics for groups that contain up to the largest 20% of the residuals. The table always includes K%. Up to 20%, the table includes 1%, 2%, 2.5%, 3%, 4%, 5%, 7.5%, 10%, and 15%. After 20%, the table includes multiples of 10%.

TNTYPE K

Determines the types of nodes to display in the tables when you issue TCRITERIA. The type of node depends on the value of K:

- 1: Displays the best terminal nodes (lowest values of MSE or MAD).
- 2: Displays the worst terminal nodes (highest values of MSE or MAD).
- 3: Displays the best and worst terminal nodes.

TRSTAT K

Displays response fits and error statistics for K terminal nodes. The nodes are the nodes with the best fit values unless you specify a different type of node with TNTYPE.

TCRITERIA K

Displays the criteria for classifying subjects into K terminal nodes. The nodes are the nodes with the highest rates of correct classifications unless you specify a different type of node with TNTYPE.

For example, with the subcommands TNTYPE 2; TRSTAT 4; TCRITERIA 3; Minitab displays the 4 worst nodes in the fits and errors table and the criteria for classifying the subjects in the 3 worst nodes.

Storage

SRTREES K

Stores the optimal tree as well as additional candidate trees that are smaller than the optimal tree, the optimal tree, and up to K larger than optimal trees.

SFIT C

Stores the fitted response for each record in C.

SRESID C

Stores the residual for each record in C.

SPRESID C

Stores the percent residual for each record in C.

SSTDEV C

Stores the standard deviation for each record in C.

SMSE C

Stores the mean squared error for each record in C. If the analysis uses k-fold cross-validation, then the stored value is for the terminal node when the record is in the training data set.

SMAD C

Stores the mean absolute deviation for each record in C. If the analysis uses k-fold cross-validation, then the stored value is for the terminal node when the record is in the training data set.

SMAPE C

Stores the mean absolute percent error for each record in C. If the analysis uses k-fold cross-validation, then the stored value is for the terminal node when the record is in the training data set.

STNID C

Stores the ID number of the terminal node for each row. The terminal node is also called a cluster.

IERTREE: Session command for selecting an alternative regression tree

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

IERTREE K K

Selects a different tree from the optimal one inside a tree object generated by the RTREE command.

The first argument, K, is a positive integer, that represents the regression tree object ID. The first regression tree that you create in a project has the object ID 1. The second regression tree that you create in a project with the RTREE command has the object ID 2. Alternative trees are always made from an optimal tree, so alternative trees do not have a tree object ID.

The second argument, K, is an integer that indexes a specific tree within the tree object. K values may be negative or positive values, depending on whether the alternative tree is smaller than or larger than the optimal tree. The tree object name uses:

- 0: Represents the optimal tree.
- Negative integers: Represents trees that are smaller than the optimal tree. For example, K = -4 would store the tree that is 4 smaller than the optimal.
- Positive integers: Represents trees that are larger than the optimal tree. For example, K = 2 would store the tree that is 2 larger than the optimal.

The tables and graphs for the new tree are the same as for the original tree. The new tables and graphs are in a new output tab. The same columns that the original analysis stored are also stored for the new tree. For example, if the original analysis stored the terminal node IDs in TNODEID[0] and you select an alternative tree that is 4 trees smaller than the optimal tree, then the new analysis stores the terminal node IDs in the column TNODEID[-4].

RTPREDICT: Session command for predicting responses for new observations for a regression tree

RTPREDICT K K [K]

Predicts the response values for new observations with the tree that K and K indicate.

The first argument, K, is the identification number of a tree. This K is always a positive integer. The first tree that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a tree remains the same even if you delete the output for another tree from the Navigator. For example, if you delete the output for the first tree, the second tree that you made remains tree 2.

The second argument, K, is an integer that indexes a specific tree within the tree object. K values may be negative or positive values, depending on whether the alternative tree is smaller or larger than the optimal tree. The second K must match one of the additional trees stored by the `RTREE` command.

In the tree object:

- 0 represents the optimal tree.
- Negative integers represent trees that are smaller than the optimal tree. For example, K = –4 stores the tree that is 4 smaller than the optimal.
- Positive integers represent trees that are larger than the optimal tree. For example, K = 2 stores the tree that is 2 larger than the optimal.

The third argument, K, allows you to optionally specify which output tab contains the prediction results. K is a positive integer from 1...N. K increases every time another tree with that same pruning increments. If you do not specify this argument, Minitab appends the prediction results to the last pruning.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the tree object.

If a predictor is categorical, the values in C must match the values in the tree object.

The order of the predictors follows a strict format.

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the RTREE command.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the RTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The RTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the tree.

If a predictor is categorical, the corresponding value for K must match the values in the tree object.

The order of the values follows a strict format:

1. Enter values for continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the RTREE command.
2. Enter values for categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the RTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3

Predictor	Value
C7	Urgent
C8	New

The RTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the values in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Results

TPREDICTION

Displays the prediction table.

Storage

STNID C

Stores the terminal node (cluster) ID for each new observation.

SPCLASS C

Stores the predicted response values.

CART Classification

CTREE: Session command for creating a classification tree

CTREE

Creates an optimal classification tree and a set of close-to-optimal classification trees. This analysis also provides an important variables chart, gain and lift charts, and other diagnostic measures so that you can assess the validity of your classification model.

The response variable column must be categorical. The values can be text or numeric.

If the response column has only 2 distinct values (binary response), you can specify which value is the reference event with EVENT. If you do not specify the event or if the response has more than 2 values (multinomial response), then the reference event depends on whether the column contains numbers or text. When the values are numbers, the highest value is the reference event. When the values are text, the last value alphabetically is the event. For example, if you enter the response variable as success and failure, success is the event.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable or the weights column.

RESPONSE C

Enter a column that contains 2 or more distinct values of a categorical response. C can be text or numeric.

EVENT K

Specifies the event value for a binary response. K must be one of the two response values.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Prior/Cost

Use one of the following subcommands to set the prior probabilities. If no subcommand from this set is issued, EPRIOR is the default.

- EPRIOR: Sets equal prior probabilities for all response levels.
- FPRIOR: Sets the prior probability for the jth response class to be the sample proportion of the class from both training and test data sets.
- PRIOR K...K: Specifies prior probabilities for all J response levels. The number of Ks must match the number of distinct response values. All Ks must be between 0 and 1, and their sum must equal 1.

COST K K K K K K...K K K

Specify costs for misclassification. Without this subcommand, equal costs of 1 are assumed.
Suppose the response has two levels, 1 and 2, and the syntax is COST 2 1 2.45 1 2 1.37

- The cost for assigning level 2 to an actual response level 1 is 2.45.
- The cost for assigning level 1 to an actual response level 2 is 1.37.

Note The order of the two cost sets does not matter. Also, you can specify the cost of some sets and not other sets. Minitab assumes the cost for unspecified sets is 1.

For a response with J levels, the total possible number of cost sets is J x (J - 1). You can specify a subset of the sets. The order of the three number sets does not matter, as long as the first two values in each set make a distinct pair, and the third value K is greater than 0.

Validation

Use one of the following subcommands to specify the validation method.

- CVTEST
- TEST
- NTEST

If no validation subcommand is issued, then Minitab uses the following defaults, depending on the number of rows:

The default when the number of rows ≤ 5000

CVTEST 10;

The default when the number of rows > 5000

TEST 0.3;

CVTEST K

Use K-fold cross-validation as a test method. Specify K number of folds.

K must be an integer between 2 and the sample size; 2 ≤ K ≤ 1000.

RANDOMIZE K

Specify the base number for randomly splitting the training data set into K subsets for cross-validation. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to reproduce the default results from the graphical user interace, where 12345 is the default seed. Without the RANDOMIZE subcommand, the analysis uses a randomly selected seed.

CVTEST C

Use K-fold cross-validation as a test method. Enter C to specify the rows for each fold.

C can be numeric or text and must be the same length as the response column.

TEST K

Specify a proportion of rows for the test set. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to reproduce the default results from the graphical user interace, where 12345 is the default seed. Without the RANDOMIZE subcommand, the analysis uses a randomly selected seed.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

NTEST

No model validation.

Use `SSAMPLE C` to store a column that shows the assignments for each row in the validation. When `CVTEST` is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When `TEST` is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the two values.

Options

Use one of the following subcommands to specify the node splitting method.

- `GINI`: Use the Gini method to split a node. This is the default method.
- `ENTROPY`: Use the Entropy method to split a node.
- `TWOING`: Use the Twoing method to split a node.

- CPROB: Use the class probability method to split a node.

The following subcommands are available to select an optimal tree.

- MINCOST: Use minimum cost to select the optimal tree.
- STERROR K : Use within K standard error of minimum cost to select the optimal tree.

If you do not specify the node splitting method, then the analysis uses GINI and STERROR 1.

The following subcommands are available to control tree growth due to sample size or node complexity.

- NOSPLIT K: Do not split a node if the node size is less than K. K=10 is the default.
- NOTERMINAL K: Do not create a terminal node that is smaller than K. K=3 is the default.
- MAXTDEPTH K: Set maximum tree depth to the Kth layer.
- NLIMIT K: Set the maximum number of terminal nodes to K. K=1024 is the default.

SURROGATE K

Specify K surrogates to consider when a specific predictor that splits a node has missing values.

K must be an integer between 0 and the number of predictors – 1, for example:

- K = 0: Specifies no consideration of surrogate variables.
- K = Number of total predictors – 1: Specifies to consider all predictors other than the splitter variable.

If this subcommand is not issued, K=10 is the default value.

WEIGHT C

Specify the case weights in C. The column length must be the same length as the target column length.

Values must be ≥ 0. Assigning a weight of 0 or missing excludes the row from the analysis.

MPENALTY K

Specify a penalty value for a predictor with missing values.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty.
- K = 2: Specifies highest penalty.

.

CPENALTY K

Specify a penalty value for categorical predictors that have many values.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty.
- K = 5: Specifies highest penalty.

Graphs

GTDIAG

Displays the optimal tree diagram.

GCOST

Displays the relative cost vs. number of terminal nodes plot.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GROC

Displays the ROC curve.

GGAIN

Displays the gain chart.

GLIFT K

Displays the lift chart.

The type of chart depends on the value of K. When K =

- 1: Displays the cumulative lift chart.
- 2: Displays the non-cumulative lift chart.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table, which includes the model validation method, the node splitting method, and other information about the analysis.

TRESPONSE

Displays the Response information table, which includes class information, count information, prior probabilities, and other information about the response variable.

TSUMMARY

Displays the Model summary table, which includes the number of predictors, the number of important predictors, the number of nodes, misclassification costs, and other information about the model.

TCONFUSION

Displays the Confusion matrix to compare performance measurements of the classification models.

For a binary response, issue the following commands after TCONFUSION to indicate which measurements to show.

TPOSITIVE

Displays true positive rate.

FPOSITIVE

Displays false positive rate.

TNEGATIVE

Displays true negative rate.

FNEGATIVE
Displays false negative rate.

TMISSCLASS
Displays the Misclassification matrix.

TNTYPE K
Determines the types of nodes to display in the tables when you issue TEFFECTIVE and TCRITERIA. The type of node depends on the value of K:

- 1: Displays the best terminal nodes. The best nodes have the highest event probability (binary) or highest class probability (multinomial) values.
- 2: Displays the worst terminal nodes. The worst nodes have the lowest event probability (binary) or lowest class probability (multinomial) values.
- 3: Displays the best and worst terminal nodes.

TEFFECTIVE K
Displays the effectiveness of the classification table for K terminal nodes. The K terminal nodes have the highest rates of correct classifications unless you specify a different type of node with TNTYPE.

TCRITERIA K
Displays the criteria for classifying subjects into K terminal nodes. The K terminal nodes have the highest rates of correct classifications unless you specify a different type of node with TNTYPE.

For example, with the subcommands TNTYPE 2; TEFFECTIVE 4; TCRITERIA 3; Minitab displays the 4 worst nodes in the classification table and the criteria for classifying the subjects in the 3 worst nodes.

Storage

SCTREES K
Stores the optimal tree as well as additional candidate trees that are smaller than the optimal tree, the optimal tree, and up to K larger than optimal trees.

SEPROB C C or C...C
Stores the probability of each class in the response variable. Specify as many columns as the response variable has classes.

SPCLASS C
Stores the predicted class for each record in the tree.

STNID C
Stores the ID number of the terminal node for each row. The terminal node is also called a cluster.

SODDS C
For a binary response variable, stores the odds for each record in the tree.

IECTREE: Session command for selecting an alternative classification tree

IECTREE K K
Selects a different tree from the optimal one inside a tree object generated by the CTREE command.

The first argument, K, is a positive integer, that represents the classification tree object ID. The first classification tree that you create in a project has the object ID 1. The second classification tree that you create in a project with the CTREE command has the object ID 2. Alternative trees are always made from an optimal tree, so alternative trees do not have a tree object ID.

The second argument, K, is an integer that indexes a specific tree within the tree object. K values may be negative or positive values, depending on whether the alternative tree is smaller than or larger than the optimal tree. The tree object name uses:

- 0: Represents the optimal tree.
- Negative integers: Represents trees that are smaller than the optimal tree. For example, K = -4 would store the tree that is 4 smaller than the optimal.
- Positive integers: Represents trees that are larger than the optimal tree. For example, K = 2 would store the tree that is 2 larger than the optimal.

The tables and graphs for the new tree are the same as for the original tree. The new tables and graphs are in a new output tab. The same columns that the original analysis stored are also stored for the new tree. For example, if the original analysis stored the terminal node IDs in TNODEID[0] and you select an alternative tree that is 4 trees smaller than the optimal tree, then the new analysis stores the terminal node IDs in the column TNODEID[-4].

CTPREDICT: Session command for predicting responses for new observations for a classification tree

CTPREDICT K K [K]

Predicts the response values for new observations with the optimal tree.

The first argument, K, is the identification number of a tree. This K is always a positive integer. The first tree that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a tree remains the same even if you delete the output for another tree from the Navigator. For example, if you delete the output for the first tree, the second tree that you made remains tree 2.

The second argument, K, is an integer that indexes a specific tree within the tree object. K values may be negative or positive values, depending on whether the alternative tree is smaller or larger than the optimal tree. The second K must match one of the additional trees stored by `CTREE`.

In the tree object:

- 0 represents the optimal tree.
- Negative integers represent trees that are smaller than the optimal tree. For example, K = –4 stores the tree that is 4 smaller than the optimal.
- Positive integers represent trees that are larger than the optimal tree. For example, K = 2 stores the tree that is 2 larger than the optimal.

The third argument, K, allows you to optionally specify which output tab contains the prediction results. K is a positive integer from 1...N. K increases every time another tree with that same pruning increments. If you do not specify this argument, Minitab appends the prediction results to the last pruning.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the tree object.

If a predictor is categorical, the values in C must match the values in the tree object.

The order of the predictors follows a strict format.

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the CTREE command.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the CTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The CTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the tree.

If a predictor is categorical, the corresponding value for K must match the values in the tree object.

The order of the values follows a strict format:

1. Enter values for continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the CTREE command.
2. Enter values for categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the CTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The CTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the values in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Results

TPREDICTION
Displays the prediction table.

Storage

STNID C
Stores the terminal node (cluster) ID for each new observation.

SEPROB C C or C...C
For a binary response variable, this subcommand stores the probability of the even in the first column and stores the probability of the non-event in the second column.

For a multinomial response, this subcommand stores each probability of the response level in a separate column. Enter the same number of columns as the response variable has levels.

SPCLASS C
Stores the predicted class for each record in the column that you specify.

TreeNet Regression

BRTREE: Session command for creating a gradient boosted regression tree

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BRTREE
Creates an optimal gradient boosted regression tree and a set of close-to-optimal gradient boosted regression trees. This analysis also provides residual plots, an important variables chart, and other diagnostic measures so that you can assess the validity of your regression model.

Use the BOTTOM subcommand to perform automatic predictor elimination. The removal of the least important predictors lets you focus on the predictors that are vital to a good model so that you can more easily interpret the model and specify values that optimize your processes. The algorithm removes the least important predictors automatically, shows you results that let you compare models with different numbers of predictors, and produces results for the set of predictors with the maximum R^2 value.

The response variable column must be continuous and must contain at least 2 distinct numeric values.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable or the weights column.

RESPONSE C
Enter a column that contains 2 or more distinct values of a continuous response. C must be numeric.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Interaction

Use one of the following subcommands to set the predictor interaction options. If no subcommand from this set is issued, AINTERACT is the default. Minitab ignores NINTERACT and SINTERACT when you use BOTTOM or TOP from the Predictor Elimination subcommands.

AINTERACT

Allow all predictor interactions when growing the sequence of trees.

NINTERACT

Do not allow predictor interactions when growing the sequence of trees. Minitab uses only the additive model.

SINTERACT K C...C

Allow two or more selected predictor interactions up to K^{th} order when growing the sequence of trees.

K is a positive number greater than 1. A large K, such as K = 5, may be similar to the results with AINTERACT.

Predictor Elimination

Use one of the following subcommands to set the predictor elimination options. After you perform use predictor elimination, you can use the BRTMODEL command to produce results for any of the models in the Predictor Elimination table. When you use predictor elimination, Minitab ignores NINTERACT and SINTERACT from the Interaction subcommands.

BOTTOM

Specify to eliminate the least important predictors from the model first.

Use the following subsubcommands to specify details about the predictor elimination. These subsubcommands have to come after BOTTOM or TOP in the session commands that you specify.

PCOUNT K

Specify how many predictors with positive importance values to eliminate at each step. In addition to the K predictors that you specify, the algorithm eliminates any predictors with 0 importance at each step. Without this subsubcommand, the algorithm eliminates 1 predictor with positive importance at each step. If you specify a value greater than or equal to the number of predictors, the algorithm eliminates all but 1 predictor.

NSTEP K

Specify the maximum number of elimination steps. Without this subsubcommand or if K is more than the number of predictors, the algorithm continues until the number of predictors is less than or equal to the argument for PCOUNT.

CORE C...C

Specify predictors to retain in the model until the elimination of all other predictors, irrespective of importance scores. Columns that you specify have to be columns that follow the CONTINUOUS or CATEGORICAL subcommands.

BOTH

Display a Predictor Elimination table that includes R^2 values for both the training set and the test set. The default model results are still of the model with the highest test R^2 value. The training results are usually more ideal than the actual results for new data. The training results are for reference only.

TOP

Specify to eliminate the most important predictors from the model first. The TOP subcommand has the same subsubcommands as the BOTTOM subcommand.

Validation

Use one of the following subcommands to specify the validation method.

- CVTEST
- TEST
- NTEST

If no validation subcommand is issued, then Minitab uses the following defaults, depending on the number of rows:

The default when the number of rows ≤ 500

- CVTEST 5; when the subcommands do not include TOP or BOTTOM
- CVTEST 3; when the subcommands include TOP or BOTTOM

The default when the number of rows > 500

TEST 0.3;

RANDOMIZE 12345;

CVTEST K

Use K-fold cross-validation as a test method. Specify K number of folds.

K must be an integer between 2 and the sample size; $2 \leq K \leq 100$.

CVTEST C

Use K-fold cross-validation as a test method. Enter C to specify the rows for each fold.

C is used to define the rows for each fold. The column must contain at least 2 distinct values. The column can be numeric or text and must be the same length as the response column.

TEST K

Specify a proportion of rows for the test set. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be a positive integer. Use 12345 to get the same results as Minitab's defaults.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

NTEST

No model validation.

Use SSAMPLE C to store a column that shows the assignments for each row in the validation. When CVTEST is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When TEST is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the two values.

Options

Use one of the following subcommands to specify the best fit method used in the analysis.

- LSE: Use the Least Squared Errors method to find the best fits. This is the default method.
- LAD: Use the Least Absolute Deviation method to find the best fits.
- HUBER K: Use the Huber loss to find the best fits. K is the switching value from squared loss to absolute loss in the hybrid Huber loss function. K must be between 0 and 1. K = 0.9 is the default value.

NTREE K

Set the number of trees to build. K must be a positive integer less than 5000. K = 300 is the default value.

Choose either MAXNODE or MAXTDEPTH. If neither of the following subcommands are issued, then the analysis uses MAXNODE 6.

MAXNODE K

Sets the maximum terminal nodes per tree. K must be a positive integer between 2 and 2000.

MAXTDEPTH K

Sets the maximum tree depth to the K^{th} layer. K must be a positive integer between 2 and 1000. MAXTDEPTH tends to give balanced trees.

NOTERMINAL K

Do not create terminal nodes smaller than K. K must be a positive integer greater than or equal to 3; K = 3 is the default value.

LRATE K

Specify the learning rate of the boosting procedure to protect from overfitting. K must be from 0.0001 to 1.

K = 0.01 is the default if the number of cases in your training data is 1000 or less. For data sets with more than 1000 cases, the default learning rate is is max[0.01, 0.1 * min(1.0, N/10000)]. For example, when the data set has 9000 responses, then the learning rate = 0.09.

SUBSAMPLE K

Specify the fraction of subsampling to use at each iteration to build a tree to protect from overfitting. K must be between 0 and 1; K = 0.5 is the default.

NPRED K

Specify the number of predictors randomly selected to come up with each node splitting rule in a tree. K must be an integer between 0 and 100; all predictor variables are used by default. When K =

- 0: The number of predictors used becomes SQRT(total number of predictors) rounded down to the nearest integer.
- 100: All predictor variables are used.

- 0 < K < 100: The number of predictor variables used becomes (total number of predictors) x (K / 100) rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors used becomes 1.

BRANDOMIZE K

Specify the base number to randomly generate subsamples to build a tree. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

MPENALTY K

Specify a penalty value for a predictor with missing values.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty (default).
- K = 2: Specifies highest penalty.

.

CPENALTY K

Specify a penalty value for categorical predictors that have many values.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty (default).
- K = 5: Specifies highest penalty.

WEIGHT C

Specifies the weights in column C. Values in C must be ≥ 0, but rows with 0 values are unused in the analysis. Length of C must be the same as the response column length.

Graphs

GRSQ

Displays the R^2 vs number of trees graph if `LSE` or `HUBER` loss is used. If `LAD` is used, `GRSQ` displays the mean absolute deviation (MAD) vs number of trees graph.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GFIT

Displays the fits vs actual response values plot.

GRBOX

Displays the boxplot of the residuals. If the analysis uses a validation method, the graph has separate boxplots for the training data set and the test data set.

RTYPE K

Specifies the type of residual to plot. When K =

- 1: Displays regular residuals.
- 2: Displays percent residuals.

GOPART K K

Displays one predictor partial dependence plots for the top K important predictors.

The first K must be a positive integer less than the total number of predictors. K indicates the number of important predictors.

The second K must be either 1 or 2. When K =

- 1: Center the y–values before plotting (default).
- 2: Do not center the y–values before plotting.

GTPART K K

Displays two predictor partial dependance plots for the top K important variables.

The first K must be a positive integer less than the total number of predictors. K indicates the number of important variables.

The second K must be either 1 or 2. When K =

- 1: Center the y–values before plotting (default).
- 2: Do not center the y–values before plotting.

Use the following subcommands to specify the types of plots to display. If neither subcommand is specified, the 3D surface plot is displayed.

- `SURFACE` : Displays a 3D surface plot.
- `GCONTOUR` : Displays a contour plot.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table, which includes the model validation method, the criterion for selecting the optimal number of trees, and other information about the analysis.

TRESPONSE

Displays the Response information table, which includes descriptive statistics about the response variable.

TSUMMARY

Displays the Model summary table, which includes the number of predictors, the number of important predictors, the number of trees grown, the R^2 statistic, and other information about the model.

TINTERACT K

Displays the 2-way interaction strength tables. K must be 0, 1, or 2. When K =

- 0: Displays the % of total squared error explained by a 2-way interaction.
- 1: Displays the % of squared error due to specific predictor pairs.

- 2: Displays both the % of total squared error and the % of squared error.

Only the top 5 predictors are displayed in the tables.

TPERROR K

Displays percent of error statistics due to largest residuals. K represents the percent of largest residuals being displayed.

0 < K ≤ 100. K = 20 is the default, which means the table includes statistics for groups that contain up to the largest 20% of the residuals. The table always includes K%. Up to 20%, the table includes 1%, 2%, 2.5%, 3%, 4%, 5%, 7.5%, 10%, and 15%. After 20%, the table includes multiples of 10%.

Storage

SBRTREE

Stores the boosted regression tree object for prediction.

SFIT C

Stores the fitted response for each record in C.

SRESID C

Stores the residual for each record in C.

SPRESID C

Stores the percent residual for each record in C.

BRTMODEL: Session command for Hyperparameter Tuning and for selecting an alternative **TreeNet® Regression** model from **Discover Key Predictors**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BRTMODEL K

After you create a TreeNet® Regression model with BRTREE, BRTMODEL completes one of two follow-up analyses.

- With the PEMODEL subcommand, BRTMODEL produces results for a model from among the models from the predictor elimination algorithm.
- With one or more of the LRATE, SUBSAMPLE, MAXNODE, MAXTDEPTH, NTREE, NPRED, and FULLCOMB subcommands, BRTMODEL tunes the hyperparameters for a TreeNet Regression model to find the combination that produces the best model.

K is a positive integer that represents a TreeNet® Regression object ID. The first TreeNet® Regression that you create in a project has the object ID 1. The second TreeNet® Regression model that you create with the BRTREE command has the object ID 2, and so on. You can use BRTMODEL with the PEMODEL subcommand only for instances of BRTREE that use the TOP subcommand, which indicates that the TreeNet® Regression model used the predictor elimination algorithm to exclude the least important predictors. You can use BRTMODEL with the subcommands for hyperparameter tuning with any TreeNet® Regression model. The subcommands for predictor elimination and for hyperparameter tuning are mutually exclusive.

Minitab recreates the same tables and graphs for the new model as for the original object ID that you specify. The tables and graphs for the new model are in a new output tab. The same columns that the original analysis stored are also stored for the new model. The storage for the new model is in a new column. For example, if the

original analysis stored the fitted values in a column titled "Fit," then the new analysis also titles an empty column "Fit_1" and stores the fitted values.

For purposes of using a command like BRTMODEL or BRPREDICT, the TreeNet® Regression object ID that comes from the BRTMODEL command is the next TreeNet® Regression object ID. For example, if you run BRTREE, the results from that model are the first TreeNet® Regression object ID: 1. If you then use BRTMODEL to select a different model from the results, that model is the second TreeNet® Regression object ID: 2. The ARTMODEL command with the TN subcommand also creates a TreeNet® Regression object ID.

Predictor elimination

PEMODEL K

Produces TreeNet Regression results for a model from the predictor elimination steps. Select K from the numbers in the Model column of the predictor elimination table to produce results for the model in that row.

Hyperparameter tuning

When you use any of the hyperparameter tuning subcommands, Minitab produces a table that compares models with the hyperparameters that you specify. Minitab also produces the full results for the TreeNet® model with the best value of the accuracy criterion, such as the highest R^2 value.

If you do not use one of the subcommands, then the value for the hyperparameter comes from the model that BRTMODEL specifies.

If you specify more than one subcommand with parameters, then the models in the comparison table depend on whether you enter the FULLCOMB subcommand.

- If you enter FULLCOMB, then the algorithm evaluates every combination of the hyperparameters.
- If you do not enter the FULLCOMB subcommand, then the algorithm evaluates the hyperparameters from the subcommands in this order:
  1. LRATE
  2. SUBSAMPLE
  3. MAXNODE or MAXTDEPTH

For a subcommand that the algorithm has not evaluated, the algorithm uses the first value to evaluate earlier subcommands. For a subcommand that the algorithm has evaluated, the algorithm uses the best value.

For example, suppose that the algorithm receives the following subcommands:

```
LRATE 0.001 0.01 0.1;
SUBSAMPLE 0.4 0.5 0.7;
MAXNODE 4 6.
```

1. The algorithm sets the subsample proportion to 0.4 and the maximum number of terminal nodes to 4. Then, the algorithm evaluates the learning rates in order from least to greatest: 0.001, 0.01, 0.1.
2. Suppose the algorithm identifies 0.01 as the best learning rate. Then the algorithm sets the learning rate to 0.01 and the maximum number of terminal nodes to 4. Then, the algorithm evaluates the subsample proportions of 0.4, 0.5, and 0.7.
3. Suppose that the algorithm identifies 0.5 as the best subsample proportion. Then the algorithm sets the learning rate to 0.01, the subsample proportion to 0.5. Then, the algorithm evaluates the maximum numbers of nodes of 4 and 6.

4. Suppose that the algorithm identifies 6 as the best maximum number of terminal nodes. Then Minitab produces the comparison table and the results for the model with learning rate = 0.01, subsample proportion 0.5, and maximum number of terminal nodes 6.

LRATE K...K

Specifies up to 10 learning rates to evaluate. You can specify values from 0.0001 to 1.

SUBSAMPLE K...K

Specifies up to 10 subsample proportions to evaluate. You can specify positive values that are less than or equal to 1.

MAXNODE K (K K)

Specifies up to 3 maximum numbers of terminal nodes to evaluate. You can specify values from 2 to 2,000. You cannot use MAXNODE and MAXTDEPTH together.

MAXTDEPTH K (K K)

Specifies up to 3 maximum tree depths to evaluate. You can specify values from 2 to 1,000. You cannot use MAXNODE and MAXTDEPTH together.

NTREE K

The maximum number of trees to build is not a hyperparameter to tune, but you can specify a different number than the model that you specify with BCTMODEL has. Specify a single value from 1 to 5,000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

FULLCOMB

Specifies that the table of models to compare includes a row for every combination of the hyperparameters that you specify with the other subsubcommands. Otherwise, the algorithm evaluates a hyperparameter at only one combination of the other hyperparameters.

BRPREDICT: Session command for predicting responses for new observations for a gradient boosted regression tree model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BRPREDICT K

Predicts the response values for new observations with the optimal gradient boosted regression tree model.

K is the identification number of a model. This K is always a positive integer. The first model that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model.

If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on if the predictor elimination algorithm removes any predictors from the model. For example predictor elimination occurs if the BRTREE command includes the BOTTOM subcommand or if the BRTMODEL command includes the PEMODEL subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict format:

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the BRTREE command.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the BRTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The BRTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the model for the TreeNet® Regression Object ID does not include all of the predictors, then omit those predictors.

1. Enter continuous predictors that are in the model first. Use the order of the predictors from the CONTINUOUS subcommand from the BRTREE command.
2. Enter categorical predictors that are in the model second. Use the order of the predictors from the CATEGORICAL subcommand from the BRTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The BRTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

The predictor elimination algorithm gave a model without C3 and C4 as predictors.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the tree.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on if the predictor elimination algorithm removes any predictors from the model. For example predictor elimination occurs if the BRTREE command includes the BOTTOM subcommand or if the BRTMODEL command includes the PEMODEL subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict format:

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from BRTREE.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from BRTREE.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The `BRTREE` command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the model for the TreeNet® Regression Object ID does not include all of the predictors, then omit those predictors.

1. Enter continuous predictors that are in the model first. Use the order of the predictors from the `CONTINUOUS` subcommand from `BRTREE`.
2. Enter categorical predictors that are in the model second. Use the order of the predictors from the `CATEGORICAL` subcommand from `BRTREE`.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The `BRTREE` command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

The predictor elimination algorithm gave a model without C3 and C4.
Enter the predictors in the following order:
KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Results

TPREDICTION
Displays the prediction table.

Storage

SFIT C
Stores the final fit of each row in the column that you specify.

TreeNet Classification

BCTREE: Session command for creating a gradient boosted classification tree

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BCTREE

Creates an optimal gradient boosted classification tree and a set of close-to-optimal gradient boosted classification trees. This analysis also provides an important variables chart, gain and lift charts, and other diagnostic measures so that you can assess the validity of your classification model.

Use the BOTTOM subcommand to perform automatic predictor elimination. The removal of the least important predictors lets you focus on the predictors that are vital to a good model so that you can more easily interpret the model and specify values that optimize your processes. The algorithm removes the least important predictors automatically, shows you results that let you compare models with different numbers of predictors, and produces results for the set of predictors with the maximum R^2 value.

The response variable column must be categorical. The values can be text or numeric.

If the response column has only 2 distinct values (binary response), you can specify which value is the reference event with EVENT. If you do not specify the event or if the response has more than 2 values (multinomial response), then the reference event depends on whether the column contains numbers or text. When the values are numbers, the highest value is the reference event. When the values are text, the last value alphabetically is the event. For example, if you enter the response variable as success and failure, success is the event.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable or the weights column.

RESPONSE C
Enter a column that contains 2 or more distinct values of a categorical response. C can be text or numeric.

EVENT K
Specifies the event value for a binary response. K must be one of the two response values.

CONTINUOUS C...C
Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C
Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Class Weights

Use the following subcommand to specify the class weights.

CWEIGHT K or K K...K K
Specifies the class weights using one of two types of numeric constants.

- Single constant, K. K can be 1 or 2 only. The value 1 indicates the response classes are not weighted. The value 2 indicates the response classes should have their weights balanced to the class size. The value 2 ensures that the sum of all the weights in each class are equal.
- Pair of constants, K...K. The first K, in the pair, is the class ID. The class ID can be in the range of values 1 to M, where M is the total number of classes in response C. The second K represents the weight for the class in the pair. K > 0. The maximum number of pairs is M. If the weight for a class is not provided, a weight of 1 is the default.

If you do not use this subcommand, the default weight is 1 for all classes.

Interaction

Use one of the following subcommands to set the predictor interaction options. If no subcommand from this set is issued, AINTERACT is the default. Minitab ignores NINTERACT and SINTERACT when you use BOTTOM or TOP from the Predictor Elimination subcommands.

AINTERACT
Allow all predictor interactions when growing the sequence of trees.

NINTERACT
Do not allow predictor interactions when growing the sequence of trees. Minitab uses only the additive model.

SINTERACT K C...C
Allow two or more selected predictor interaction up to K^{th} order when growing the sequence of trees.

K is a positive number greater than 1. A large K, such as K = 5, may be similar to the results with AINTERACT.

Predictor Elimination

Use the BOTTOM or the TOP subcommand to perform predictor elimination. After you perform predictor elimination, you can use the BCTMODEL command to produce results for any of the models in the Predictor Elimination table. When you use predictor elimination, Minitab ignores NINTERACT and SINTERACT from the Interaction subcommands.

BOTTOM

Specify to eliminate the least important predictors from the model first.

Use the following optional subsubcommands to specify details about the predictor elimination. These subsubcommands have to come after BOTTOM or TOP in the session commands that you specify.

PCOUNT K

Specify how many predictors with positive importance values to eliminate at each step. In addition to the K predictors that you specify, the algorithm eliminates any predictors with 0 importance at each step. Without this subsubcommand, the algorithm eliminates 1 predictor with positive importance at each step. If you specify a value greater than or equal to the number of predictors, the algorithm eliminates all but 1 predictor.

NSTEP K

Specify the maximum number of elimination steps. Without this subsubcommand or if K is more than the number of predictors, the algorithm continues until the number of predictors is less than or equal to the argument for PCOUNT.

CORE C...C

Specify predictors to retain in the model until the elimination of all other predictors, irrespective of importance scores. Columns that you specify have to be columns that follow the CONTINUOUS or CATEGORICAL subcommands.

BOTH

Display a Predictor Elimination table that includes loglikelihood values for both the training set and the test set. The default model results are still of the model with the highest best loglikelihood value. The training results are usually more ideal than the actual results for new data. The training results are for reference only.

TOP

Specify to eliminate the most important predictors from the model first. The TOP subcommand has the same subsubcommands as the BOTTOM subcommand.

Validation

Use one of the following subcommands to specify the validation method.

- CVTEST
- TEST
- NTEST

If no validation subcommand is issued, then Minitab uses the following defaults, depending on the number of rows:

The default when the number of rows ≤ 500

- CVTEST 5; when the subcommands do not include TOP or BOTTOM
- CVTEST 3; when the subcommands include TOP or BOTTOM

The default when the number of rows > 500

TEST 0.3;

RANDOMIZE 12345;

CVTEST K

Use K-fold cross-validation as a test method. Specify K number of folds.

K must be an integer between 2 and the sample size; $2 \leq K \leq 100$.

CVTEST C

Use K-fold cross-validation as a test method. Enter C to specify the rows for each fold.

C is used to define the rows for each fold. The column must contain at least 2 distinct values. The column can be numeric or text and must be the same length as the response column.

TEST K

Specify a proportion of rows for the test set. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be a positive integer. Use 12345 to get the same results as Minitab's defaults.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

NTEST

No model validation.

Use `SSAMPLE C` to store a column that shows the assignments for each row in the validation. When `CVTEST` is used:

- The storage column, Sample_Id, stores the fold identification in C.
- C contains values 1, 2, ..., K.

When `TEST` is used:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the two values.

Options

Use one of the following subcommands to specify the number of trees for the optimal model.

- `ROC`: Use the maximum area under the ROC curve value to select the number of trees for the optimal model. ROC is only available with a binary response and is the default for a binary response.
- `MAXLKH` : Use the maximum likelihood value to select the number of trees for the optimal model.
- `MISCLASS` : Use the minimum misclassification rate to select the number of trees for the optimal model. `MISCLASS` is the default for a multinomial response.

NTREE K

Set the number of trees to build. K must be a positive integer less than 5000. K = 300 is the default value.

Choose either `MAXNODE` or `MAXTDEPTH`. If neither of the following subcommands are issued, then the analysis uses MAXNODE 6.

MAXNODE K

Sets the maximum terminal nodes per tree. K must be a positive integer between 2 and 2000.

MAXTDEPTH K

Sets the maximum tree depth to the K^{th} layer. K must be a positive integer between 2 and 1000. `MAXTDEPTH` tends to give balanced trees.

NOTERMINAL K

Do not create terminal nodes smaller than K. K must be a positive integer greater than or equal to 3; K = 3 is the default value.

LRATE K

Specify the learning rate of the boosting procedure to protect from overfitting. K must be from 0.0001 to 1

K = 0.01 is the default if the number of cases in your training data is 1000 or less. For data sets with more than 1000 cases, the default learning rate is is max[0.01, 0.1 * min(1.0, N/10000)]. For example, when the data set has 9000 responses, then the learning rate = 0.09.

SUBSAMPLE K

Specify the fraction of subsampling to use at each iteration to build a tree to protect from overfitting. K must be between 0 and 1; K = 0.5 is the default. To specify the fraction to sample from each outcome of a binary response variable, use `STSAMPLE`.

STSAMPLE K K

Use STSAMPLE to specify the fraction to sample from each outcome of a binary response variable. The first K indicates the fraction of subsampling whose response is the event, and the second K indicates the fraction of subsampling whose response is the non-event. Define which response is the event with the `EVENT` subcommand.

NPRED K

Specify the number of predictors randomly selected to come up with each node splitting rule in a tree. K must be an integer between 0 and 100; all predictor variables are used by default. When K =

- 0: The number of predictors used becomes SQRT(total number of predictors) rounded down to the nearest integer.
- 100: All predictor variables are used.
- 0 < K < 100: The number of predictor variables used becomes (total number of predictors) x (K / 100) rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors used becomes 1.

BRANDOMIZE K

Specify the base number to randomly generate subsamples to build a tree. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

MPENALTY K

Specify a penalty value for a predictor with missing values.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty.
- K = 2: Specifies highest penalty.

.

CPENALTY K

Specify a penalty value for categorical predictors that have many values.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty.
- K = 5: Specifies highest penalty.

WEIGHT C

Specifies the weights in column C. Values in C must be ≥ 0, but rows with 0 values are unused in the analysis. Length of C must be the same as the response column length.

Graphs

GLIKELIHOOD

Displays the average negative loglikelihood vs number of trees plot if `MAXLKH` is used. Displays the area under ROC curve vs number of trees plot if `ROC` is used and the response variable is binary. Displays the misclassification rate vs number of trees plot if `MISCLASS` is used.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GROC

Displays the ROC curve for training and test samples.

GGAIN

Displays the gain chart for training and test samples.

GLIFT K

Displays the lift chart for training and test samples.

The type of chart depends on the value of K. When K =

- 1: Displays the cumulative lift chart.
- 2: Displays the non-cumulative lift chart.

GBPROB

Displays the box plot of estimated event probabilities when the response variable is binary.

GOPART K K

Displays one predictor partial dependence plots for the top K important predictors.

The first K must be a positive integer less than the total number of predictors. K indicates the number of important predictors.

The second K must be either 1 or 2. When K =

- 1: Center the y–values before plotting (default).
- 2: Do not center the y–values before plotting.

GTPART K K

Displays two predictor partial dependance plots for the top K important variables.

The first K must be a positive integer less than the total number of predictors. K indicates the number of important variables.

The second K must be either 1 or 2. When K =

- 1: Center the y–values before plotting (default).

- 2: Do not center the y–values before plotting.

Use the following subcommands to specify the types of plots to display. If neither subcommand is specified, the 3D surface plot is displayed.

- `SURFACE` : Displays a 3D surface plot.
- `GCONTOUR` : Displays a contour plot.

Results

NODEFAULT
Suppresses the display of default tables and graphs.

TMETHOD
Displays the Method table, which includes the model validation method, the criterion for selecting the optimal number of trees, and other information about the analysis.

TRESPONSE
Displays the Response information table, which includes class information, count information, and other information about the response variable.

TSUMMARY
Displays the Model summary table, which includes the number of predictors, the number of important predictors, the number of trees grown, misclassification rates, and other information about the model.

TINTERACT K
Displays the 2-way interaction strength tables. K must be 0, 1, or 2. When K =

- 0: Displays the % of total squared error explained by a 2-way interaction.
- 1: Displays the % of squared error due to specific predictor pairs.
- 2: Displays both the % of total squared error and the % of squared error.

Only the top 5 predictors are displayed in the tables.

TCONFUSION
Displays the confusion matrix table to compare performance measurements of the classification models.

TMISSCLASS
Displays the misclassification table.

CTHRESHOLD K
For a binary response, K specifies the threshold value such that if the probability of the event > K, then assign the event level as the predicted response. Otherwise, assign the non-event level as the predicted response.

K must be between 0 and 1; K = 0.5 is the default. K = 0 indicates that the threshold equals the estimated event probability of the data when no model is built.

Storage

SBCTREE
Stores the boosted classification tree object for prediction.

SEPROB C C or C...C

For a binary response variable, stores the probability of event in the first column C, and stores the probability of non-event in the second column C.

For a multinomial response variable, stores each probability of response = level in columns C...C.

SPCLASS C

Stores the predicted class for each record in the tree in column C.

SODDS C

For a binary response variable, stores the odds for each record in the tree in column C.

SFIT C or C...C

For a binary response variable, stores the final fit of each row in column C.

For a multinomial response variable, stores the final fit of each row in columns C...C.

BCTMODEL: Session command for Hyperparameter Tuing and selecting an alternative **TreeNet® Classification** model from **Discover Key Predictors**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BCTMODEL K

After you create a TreeNet® Classification model with BCTREE, BCTMODEL completes one of two follow-up analyses.

- With the PEMODEL subcommand, BCTMODEL produces results for a model from among the models from the predictor elimination table.
- With one or more of the LRATE, SUBSAMPLE, MAXNODE, MAXTDEPTH, NTREE, NPRED and FULLCOMB subcommands, BCTMODEL tunes the hyperparameters for a TreeNet classification model to find the combination that produces the best model.

K is a positive integer that represents a TreeNet® Classification object ID. The first TreeNet® Classification that you create in a project has the object ID 1. The second TreeNet® Classification model that you create with the BCTREE command has the object ID 2, and so on. You can use BCTMODEL with the PEMODEL subcommand only for instances of BCTREE that use the TOP subcommand, which indicates that the TreeNet® Classification model used the predictor elimination algorithm to exclude the least important predictors. You can use BCTMODEL with the subcommands for hyperparameter tuning with any TreeNet® Classification model. The subcommands for predictor elimination and for hyperparameter tuning are mutually exclusive.

Minitab recreates the same tables and graphs for the new model as for the original object ID that you specify. The tables and graphs for the new model are in a new output tab. The same columns that the original analysis stored are also stored for the new model. The storage for the new model is in a new column. For example, if the original analysis stored the fitted values in a column titled "Fit," then the new analysis also titles an empty column "Fit_1" and stores the fitted values.

For purposes of using the a command like BCTMODEL or BCPREDICT, the TreeNet® Classification object ID that comes from the BCTMODEL is the next TreeNet® Classification object ID. For example, if you run BCTREE, the results from that model are the first TreeNet® Classification object ID: 1. If you then use BCTMODEL to select a different model from the Discover Best Model results, that model is the second TreeNet® Classification object ID: 2. The ACTMODEL command with the TN subcommand also creates a TreeNet® Regression object ID.

Predictor elimination

PEMODEL K

Produces TreeNet® Classification results for a model from the predictor elimination steps. Select K from the numbers in the Model column of the predictor elimination table to produce results for the model in that row.

Hyperparameter tuning

When you use any of the hyperparameter tuning subcommands, Minitab produces a table that compares models with the hyperparameters that you specify. Minitab also produces the full results for the TreeNet® model with the best value of the accuracy criterion, such as the minimum average –loglikelihood.

If you do not use one of the subcommands, then the value for the hyperparameter comes from the model that `BCTMODEL` specifies.

If you specify more than one subcommand with parameters, then the models in the comparison table depend on whether you enter the `FULLCOMB` subcommand.

- If you enter `FULLCOMB`, then the algorithm evaluates every combination of the hyperparameters.
- If you do not enter the `FULLCOMB` subcommand, then the algorithm evaluates the hyperparameters from the subcommands in this order:
  1. `LRATE`
  2. `SUBSAMPLE`
  3. `MAXNODE` or `MAXTDEPTH`

For a subcommand that the algorithm has not evaluated, the algorithm uses the smallest value to evaluate earlier subcommands. For a subcommand that the algorithm has evaluated, the algorithm uses the best value.

For example, suppose that the algorithm receives the following subcommands:

```
LRATE 0.001 0.01 0.1;
SUBSAMPLE 0.4 0.5 0.7;
MAXNODE 4 6.
```

1. The algorithm sets the subsample proportion to 0.4 and the maximum number of terminal nodes to 4. Then, the algorithm evaluates the learning rates in order from least to greatest: 0.001, 0.01, 0.1.
2. Suppose the algorithm identifies 0.01 as the best learning rate. Then the algorithm sets the learning rate to 0.01 and the maximum number of terminal nodes to 4. Then, the algorithm evaluates the subsample proportions of 0.4, 0.5, and 0.7.
3. Suppose that the algorithm identifies 0.5 as the best subsample proportion. Then the algorithm sets the learning rate to 0.01, the subsample proportion to 0.5. Then, the algorithm evaluates the maximum numbers of nodes of 4 and 6.
4. Suppose that the algorithm identifies 6 as the best maximum number of terminal nodes. Then Minitab produces the comparison table and the results for the model with learning rate = 0.01, subsample proportion 0.5, and maximum number of terminal nodes 6.

LRATE K...K

Specifies up to 10 learning rates to evaluate. You can specify values from 0.0001 to 1.

SUBSAMPLE K...K

Specifies up to 10 subsample proportions to evaluate. You can specify positive values that are less than or equal to 1.

MAXNODE K (K K)

Specifies up to 3 maximum numbers of terminal nodes to evaluate. You can specify values from 2 to 2,000. You cannot use MAXNODE and MAXTDEPTH together.

MAXTDEPTH K (K K)

Specifies up to 3 maximum tree depths to evaluate. You can specify values from 2 to 1,000. You cannot use MAXNODE and MAXTDEPTH together.

NTREE K

The maximum number of trees to build is not a hyperparameter to tune, but you can specify a different number from model that you specify with BCTMODEL. Specify a single value from 1 to 5,000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

FULLCOMB

Specifies that the table of models to compare includes a row for every combination of the hyperparameters that you specify with the other subsubcommands. Otherwise, the algorithm evaluates a hyperparameter at only one combination of the other hyperparameters.

BCPREDICT: Session command for predicting responses for new observations with a gradient boosted classification tree model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

BCPREDICT K

Predicts the response values for new observations with the optimal gradient boosted classification tree model.

K is the identification number of a model. This K is always a positive integer. The first model that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model.

If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on if the predictor elimination algorithm removes any predictors from the model. For example predictor elimination occurs if the BCTREE command includes the BOTTOM subcommand or if the BCTMODEL command includes the PEMODEL subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict format:

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the BCTREE command.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the BCTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The BCTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the model for the TreeNet® Classification Object ID does not include all of the predictors, then omit those predictors.

1. Enter continuous predictors that are in the model first. Use the order of the predictors from the CONTINUOUS subcommand from the BCTREE command.
2. Enter categorical predictors that are in the model second. Use the order of the predictors from the CATEGORICAL subcommand from the BCTREE command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The BCTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

The predictor elimination algorithm gave a model without C3 and C4 as predictors.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. KPREDICTORS can be repeated up to three times.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on if the predictor elimination algorithm removes any predictors from the model. For example predictor elimination occurs if the BCTREE command includes the BOTTOM subcommand or if the BCTMODEL command includes the PEMODEL subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict format:

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from BCTREE.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from BCTREE.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100

Predictor	Value
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The BCTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the model for the TreeNet® Classification Object ID does not include all of the predictors, then omit those predictors.

1. Enter continuous predictors that are in the model first. Use the order of the predictors from the CONTINUOUS subcommand from BCTREE.
2. Enter categorical predictors that are in the model second. Use the order of the predictors from the CATEGORICAL subcommand from BCTREE.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The BCTREE command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

The predictor elimination algorithm gave a model without C3 and C4.

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Results

TPREDICTION

Displays the prediction table.

Storage

SEPROB C C or C...C

For a binary response variable, this subcommand stores the probability of the even in the first column and stores the probability of the non-event in the second column.

For a multinomial response, this subcommand stores each probability of the response level in a separate column. Enter the same number of columns as the response variable has levels.

SPCLASS C

Stores the predicted class for each record in the column that you specify.

SFIT C or C...C

Stores the final fit of each row in column C for a binary response variable.

Stores the final fit of each row in columns C...C for a multinomial response variable.

Random Forests Regression

RRFOREST: Session command for creating a random forest of regression trees

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

RRFOREST

Creates a random forest of regression trees using bootstrap samples from the original data. Based on the results, one can gain insights into the relationship between a continuous response and the important predictors out of many candidate predictors, and predict response values for new observations with great accuracy. This analysis also provides residual plots, an important variables chart, and other diagnostic measures so that you can assess the validity of your regression model.

The response variable column must be continuous and must contain at least 2 distinct numeric values.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable.

RESPONSE C

Enter a column that contains 2 or more distinct values of a continuous response. C must be numeric.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Validation

Use one of the following subcommands to specify the validation method.

- OOB
- TEST

If no validation subcommand is issued, then Minitab uses OOB.

OOB

Use out-of-bag data for validation. RRFOREST uses bootstrap sampling for every tree. Every record in a bootstrap sample is randomly selected with replacement from the original data set. Out of the total observations of the original data set, some records will be left out from every bootstrap sample. The left-out rows from a bootstrap sample make an out-of-bag data set.

TEST K

Specify a proportion of rows for the test set if you only enter a constant. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to get the same results as Minitab's defaults.

SSAMPLE C

Store a column that shows the assignments for each row in the validation. The storage column, Sample_Id, stores the training and test sample assignments in C. C is a binary text column that contains Training and Test as the two values.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

Options

NTREE K (K)

Set K bootstrap samples to grow K trees. The first constant specifies the number of bootstrap samples and the number of trees produced by the analysis. You must enter a value between 3 and 3000.

The second optional constant sets the bootstrap sample size. You must enter a value greater than or equal to 5. If you don't enter a value or if you enter a value that is greater than the training data size, Minitab uses a sample size equal to the training data size.

If you don't specify NTREE, Minitab selects 300 bootstrap samples with a sample size equal to the training data size.

NPRED K

Specify the number of predictors randomly selected to come up with each node splitting rule in a tree. K must be an integer between -1 and 100. When K =

- -1: The number of predictors used becomes SQRT(total number of predictors) rounded down to the nearest integer.
- 0: All predictor variables are used.
- 0 < K < 100: The number of predictor variables used becomes (total number of predictors) x (K / 100) rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors used becomes 1.

If you don't specify NPRED, K = -1, which uses SQRT(total number of predictors), is the default.

BRANDOMIZE K

Specify the base number to randomly generate the bootstrap samples. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

NOSPLIT K

Do not split a node if the node size is less than K. K=2 is the default.

MPENALTY K

Specify a penalty value for a predictor with missing values. The default value is K = 0 if you do not specify this subcommand.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty.
- K = 2: Specifies highest penalty.

CPENALTY K

Specify a penalty value for categorical predictors that have many values. The default value is K = 0 if you do not specify this subcommand.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty.
- K = 5: Specifies highest penalty.

Graphs

GRSQ

Displays the R^2 vs number of trees plot.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GFIT

Displays fits vs actual response values plot.

GRBOX

Displays boxplot of the residuals.

RTYPE K

Specifies the type of residual to plot. When K =

- 1: Displays regular residuals.
- 2: Displays percent residuals.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table, which includes information about model validation, number of bootstrap samples, and other information about the analysis.

TRESPONSE

Displays the Response information table, which includes descriptive statistics about the response variable.

TSUMMARY

Displays the Model summary table, which includes the statistics that describe the model. Includes the number of predictors, important predictors, and goodness-of-fit information about the model.

TPERROR K

Displays percent of error statistics due to largest residuals. K represents the percent of largest residuals being displayed.

$0 < K \leq 100$. K = 20 is the default, which means the table includes statistics for groups that contain up to the largest 20% of the residuals. The table always includes K%. Up to 20%, the table includes 1%, 2%, 2.5%, 3%, 4%, 5%, 7.5%, 10%, and 15%. After 20%, the table includes multiples of 10%.

Storage

SRRTREE K

Stores the model object for prediction. You cannot predict from the model unless you store the model object. K specifies the number of trees for prediction. The default value for K is 300.

SFIT C

Stores the fitted response for each record in C.

SRESID C

Stores the residual for each record in C.

SPRESID C

Stores the percent residual for each record in C.

RRPREDICT: Session command for predicting responses for new observations for a random forest regression model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

RRPREDICT K

Predicts the response values for new observations with the optimal model.

K is the identification number of a model object. This K is always a positive integer. The first model that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model. If a predictor is categorical, the values in C must match the values in the model.

The order of the predictors follows a strict format.

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from RRFOREST.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from RRFOREST.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The RRFOREST command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. KPREDICTORS can be repeated up to three times.

If a predictor is categorical, the corresponding value for K must match the values in the model.

The order of the values follows a strict format:

1. Enter values for continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the RRFOREST command.
2. Enter values for categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the RRFOREST command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A

Predictor	Value
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The RRFOREST command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the values in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Results

TPREDICTION

Displays the prediction table.

Storage

SFIT C

Stores the fit of each row in the column that you specify.

Random Forests Classification

RCFOREST: Session command for creating a random forest of classification trees

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

RCFOREST

Creates a random forest of classification trees using bootstrap samples from the original data. This analysis also provides an important variables chart, gain and lift charts, and other diagnostic measures so that you can assess the validity of your classification model.

The response variable column must be categorical. The values can be text or numeric.

If the response column has only 2 distinct values (binary response), you can specify which value is the reference event with EVENT. If you do not specify the event or if the response has more than 2 values (multinomial response), then the reference event depends on whether the column contains numbers or text. When the values are numbers, the highest value is the reference event. When the values are text, the value order of the column is used to define the event class for the binary case and the order of displaying classes for the multinomial case. For example, if you enter the response variable as success and failure, success is the event.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable.

RESPONSE C

Enter a column that contains 2 or more distinct values of a categorical response. C can be text or numeric.

EVENT K

Specifies the event value for a binary response. K must be one of the two response values.

CONTINUOUS C...C

Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Class Weights

Use the following subcommand to specify the class weights.

CWEIGHT K or K K...K K

Specifies the class weights using one of two types of numeric constants.

- Single constant, K. K can be 1 or 2 only. The value 1 indicates the response classes are not weighted. The value 2 indicates the response classes are weighted to ensure an equal sample size across all classes.
- Pair of constants, K...K. The first K, in the pair, is the class ID. The class ID can be in the range of values 1 to M, where M is the total number of classes in response C. The second K represents the weight for the class in the pair. K > 0. The maximum number of pairs is M. If the weight for a class is not provided, a weight of 1 is the default.

If you do not use this subcommand, the default weight is 1 for all classes.

Validation

Use one of the following subcommands to specify the validation method.

- OOB
- TEST

If no validation subcommand is issued, then Minitab uses OOB.

OOB

Use out-of-bag data for validation. RCFOREST uses bootstrap sampling for every tree. Every record in a bootstrap sample is randomly selected with replacement from the original data set. Out of the total observations of the original data set, some records will be left out from every bootstrap sample. The left-out rows from a bootstrap sample make an out-of-bag data set.

TEST K

Specify a proportion of rows for the test set if you only enter a constant. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be positive integer. Use 12345 to get the same results as Minitab's defaults.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

SSAMPLE C

Store a column that shows the assignments for each row in the validation. The storage column, Sample_Id, stores the training and test sample assignments in C. C is a binary text column that contains Training and Test as the two values.

Options

NTREE K (K)

Set K bootstrap samples to grow K trees. The first constant specifies the number of bootstrap samples and the number of trees produced by the analysis. You must enter a value between 3 and 3000.

The second optional constant sets the bootstrap sample size. You must enter a value greater than or equal to 5. If you don't enter a value or if you enter a value that is greater than the training data size, Minitab uses a sample size equal to the training data size.

If you don't specify `NTREE`, Minitab selects 300 bootstrap samples with a sample size equal to the training data size.

NPRED K

Specify the number of predictors randomly selected to come up with each node splitting rule in a tree. K must be an integer between -1 and 100. When K =

- -1: The number of predictors used becomes SQRT(total number of predictors) rounded down to the nearest integer.
- 0: All predictor variables are used.
- 0 < K < 100: The number of predictor variables used becomes (total number of predictors) x (K / 100) rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors used becomes 1.

If you don't specify `NPRED`, K = -1, which uses SQRT(total number of predictors), is the default.

BRANDOMIZE K

Specify the base number to randomly generate the bootstrap samples. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

NOSPLIT K

Do not split a node if the node size is less than K. K=2 is the default.

MPENALTY K

Specify a penalty value for a predictor with missing values. The default value is K = 0 if you do not specify this subcommand.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty.
- K = 2: Specifies highest penalty.

CPENALTY K

Specify a penalty value for categorical predictors that have many values. The default value is K = 0 if you do not specify this subcommand.

0.0 ≤ K ≤ 5.0, for example:

- K = 0: Specifies no penalty.
- K = 5: Specifies highest penalty.

Graphs

GMISCLASSIFICATION

Displays the misclassification rate vs number of trees plot.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables.
- 0: Displays all predictors.

GINI

Specifies the Gini method to calculate the relative importance scores for predictors.

PERMUTE

Specifies the permutation method to calculate the relative importance scores for predictors.

If you don't specify either `GINI` or `PERMUTE`, Minitab uses `PERMUTE` when the number of observations is less than or equal to 5000. When the number of observations is greater than 5000, Minitab uses `GINI`.

GROC

Displays the ROC curve.

GGAIN

Displays the gain chart.

GLIFT K

Displays the lift chart.

The type of chart depends on the value of K. When K =

- 1: Displays the cumulative lift chart.
- 2: Displays the non-cumulative lift chart.

Results

NODEFAULT
Suppresses the display of default tables and graphs.

TMETHOD
Displays the Method table, which includes information about model validation, number of bootstrap samples, and other information about the analysis.

TRESPONSE
Displays the Response table, which includes class information, count information, and other information about the responses.

TSUMMARY
Displays the Summary table, which includes the statistics that describe the model. Includes the number of predictors, important predictors, misclassification rate, and goodness-of-fit information about the model.

TCONFUSION
Displays the confusion matrix table to compare performance measurements of the classification model.

TMISSCLASS
Displays the misclassification table.

Storage

SCTREE
Stores the model object for prediction. You cannot predict from the model unless you store the model object.

SEPROB C C or C...C
For a binary response variable, stores the probability of event in the first column C, and the probability of non-event in the second column C.

For a multinomial response variable with K levels, stores each probability of response = level in columns C...C respectively.

SPCLASS C
Stores the predicted class for each record in the training data set in column C.

SODDS C
For a binary response variable, stores the odds for each record in the training data set in column C.

RCPREDICT: Session command for predicting responses for new observations with a random classification forest model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

RCPREDICT K
Predicts the response values for new observations with the optimal model.

K is the identification number of a model object. This K is always a positive integer. The first model that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model. If a predictor is categorical, the values in C must match the values in the model.

The order of the predictors follows a strict format.

1. Enter continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from RCFOREST.
2. Enter categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from RCFOREST.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The RCFOREST command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. KPREDICTORS can be repeated up to three times.

If a predictor is categorical, the corresponding value for K must match the values in the model.

The order of the values follows a strict format:

1. Enter values for continuous predictors first. Use the order of the predictors from the CONTINUOUS subcommand from the RCFOREST command.
2. Enter values for categorical predictors second. Use the order of the predictors from the CATEGORICAL subcommand from the RCFOREST command.

For example, suppose you fit a tree using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The RCFOREST command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the values in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Results

TPREDICTION
Displays the prediction table.

Storage

SEPROB C C or C...C
For a binary response variable, this subcommand stores the probability of the event class for each new record in the first column and stores the probability of the non-event class in the second column.

For a multinomial response, this subcommand stores each probability of the response level in a separate column. Enter the same number of columns as the number of distinct levels in the response column.

SPCLASS C
Stores the predicted class for each record in the column that you specify.

MARS Regression

MARS: Session command for creating a multi-adaptive regression spline model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

MARS
Creates a MARS regression model with an optimal number of basis functions and a set of other models with different numbers of basis functions. This analysis also provides an important variables chart. For additive models, this analysis also provides one-predictor, partial dependence plots that provide deeper information about the effects of the predictors.

The response variable column must be continuous and must contain at least 2 distinct numeric values.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values from the response variable or the weights column.

RESPONSE C
Enter a column that contains 2 or more distinct values of a continuous response. C must be numeric.

CONTINUOUS C...C
Specifies the continuous predictors are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C
Specifies the categorical predictors are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Interaction

Use one of the following subcommands to set the predictor interaction options. If no subcommand from this set is issued, NINTERACT is the default.

NINTERACT

Do not allow predictor interactions in the model. Minitab uses only the additive model.

SINTERACT K [C...C]

Allow predictor interactions up to the K^{th} order when building the model. Optionally, specify 2 or more columns that are arguments for CONTINUOUS and CATEGORICAL to allow only interactions among the variables that you specify.

K is a positive number. K = 1 is equivalent to the use of the NINTERACT subcommand. The maximum value of K is 4.

Validation

Use one of the following subcommands to specify the validation method.

- CVTEST
- TEST
- NTEST

If no validation subcommand is issued, then Minitab uses the following defaults, depending on the number of rows:

The default when the number of rows ≤ 500

CVTEST 5

RANDOMIZE 12345;

The default when the number of rows > 500

TEST 0.3;

RANDOMIZE 12345;

CVTEST K

Use K-fold cross-validation as a test method. Specify K number of folds.

K must be an integer of at least 2. The maximum value of K is the sample size or 100, whichever is smaller.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into folds. K sets the starting point for Minitab's random number generator.

K must be a positive integer. Use 12345 to get the same results as Minitab's defaults.

SSAMPLE C

Store a column that shows the assignments for each row in the validation. The storage column, Sample_Id, stores the fold identification in C. The column contains the values 1, 2, ..., K where K is the number of folds.

CVTEST C

Use K-fold cross-validation as a test method. Enter C to specify the rows for each fold.

C is used to define the rows for each fold. The column must contain at least 2 distinct values. The column can be numeric or text and must be the same length as the response column.

TEST K

Specify a proportion of rows for the test set. K is the proportion of the data to use for the test set. K must be between 0 and 1.

RANDOMIZE K

Specify the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for Minitab's random number generator.

K must be a positive integer. Use 12345 to get the same results as Minitab's defaults.

SSAMPLE C

Store a column that shows the assignments for each row in the validation. The storage column, Sample_Id, stores the training and test sample assignments in C. The storage column is a binary text column that contains Training and Test as the two values.

TEST C K

Use specific rows for the test set. Enter C to specify which rows are in the test set. C is an ID column that must contain only 2 values. The value that you specify with K indicates that the row is in the test set. The other value is for rows that are not in the test set.

NTEST

No model validation.

Options

Use one of the following subcommands to specify the best fit method used in the analysis.

- MAXRSQ: Use the maximum R-squared to select the optimal number of basis functions. This is the default method.
- MAD: Use the minimum mean absolute deviation to select the optimal number of basis functions.

BFUNCTIONS K

Set the maximum number of basis functions to evaluate in the construction of the model. K must be a positive integer of 200 or less. K = 30 is the default value.

MINOBS K

Specifies to use an algorithm to determine the minumum number of data points between changes in the basis functions or specifies the minumum number.

K = 0

The analysis uses sample size and model complexity to automatically select a value.

1 ≤ K ≤ N

A value of 1 indicates that consecutive data points are eligible to be points where the basis function changes. The value of 1 allows the most rapid changes in the model predictions. Use larger values to create smoother, potentially less locally-accurate models to explore more general relationships. N is the number of rows in the column for the response variable. Values larger than the sample size result in an error.

MPENALTY K

Specify a penalty value for a predictor with missing values.

0.0 ≤ K ≤ 2.0, for example:

- K = 0: Specifies no penalty (default).
- K = 2: Specifies highest penalty.

CPENALTY K

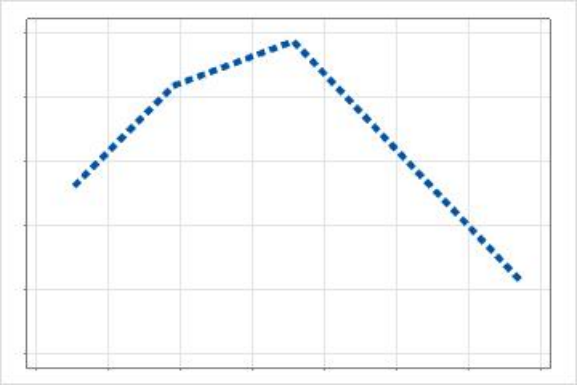
Specify a penalty value for categorical predictors that have many values.

0.0 ≤ K ≤ 5.0, for example:

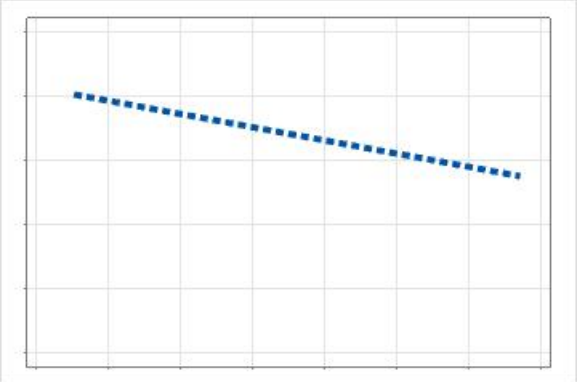
- K = 0: Specifies no penalty (default).
- K = 5: Specifies highest penalty.

NTRANSFORM C...C

Specifies predictors that have only a linear effect in the model. Another way to describe this effect is that every predictor has only 1 basis function with a knot at the minimum value of the predictor. For example, suppose a predictor in an additive model has multiple basis functions to model curvature. The partial dependence plot that results for a predictor with 3 basis functions has the following form:



Without a transformation for the predictor, the predictor has only 1 basis function. The partial dependence plot that results from 1 basis function is a straight line:



WEIGHT C

Specifies the weights in column C. Values in C must be ≥ 0, but rows with 0 values are unused in the analysis. Length of C must be the same as the response column length.

Graphs

GRSQ

Displays the R^2 vs number of basis functions graph if MAXRSQ is used. If MAD is used, GRSQ displays mean absolute deviation (MAD) vs number of basis functions graph.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables. For example, `GIMPORTANCE 50` displays 50% of the important predictors.
- 0: Displays all predictors.

GFIT

Displays the fits vs actual response values plot.

GRBOX

Displays the boxplot of the residuals. If the analysis uses a validation method, the graph has separate boxplots for the training data set and the test data set.

RTYPE K

Specifies the type of residual to plot. When K =

- 1: Displays regular residuals.
- 2: Displays percent residuals.

GPLOT2D

Displays one-predictor, partial dependence plots for additive models. The number of plots is the same as the number of variables in the final model. The analysis ignores this subcommand when the analysis includes SINTERACT.

Results

NODEFAULT

Suppresses the display of default tables and graphs.

TMETHOD

Displays the Method table, which includes the model validation method, the criterion for selecting the optimal number of basis functions, and other information about the analysis.

TRESPONSE

Displays the Response information table, which includes descriptive statistics about the response variable.

TSUMMARY

Displays the Model summary table, which includes the number of predictors, the number of important predictors, the maximum number of basis functions to evaluate, the R^2 statistic, and other information about the model.

TEQUATION

Displays the model equation. The equation has a coefficient for each basis function.

TPERROR K

Displays the percent of error statistics due to the largest residuals. K represents the percent of the largest residuals being displayed.

0 < K ≤ 100. K = 20 is the default, which means the table includes statistics for groups that contain up to the largest 20% of the residuals. The table always includes K%. Up to 20%, the table includes 1%, 2%, 2.5%, 3%, 4%, 5%, 7.5%, 10%, and 15%. After 20%, the table includes multiples of 10%.

Storage

SMARS

Stores the model object to use in subsequent analyses. If you do not store the model, then you cannot perform analyses like prediction or the selection of an alternative model,

SFIT C

Stores the fitted response for each record in C.

SRESID C

Stores the residual for each record in C.

SPRESID C

Stores the percent residual for each record in C.

SAMARS: Session command for selecting an alternative MARS model with a specific number of basis functions

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

SAMARS K K

Produce results for a model with a specific number of basis functions from the search for the optimal MARS® model.

The first argument, K, is the identification number of a model. This K is always a positive integer. The first MARS® model that you create in a project is identified as 1, the second is identified as 2, and so on.

Note The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second MARS® model that you made remains model 2.

The second argument, K, is an integer that indexes a specific model within the model object. K values are 1 more than the number of basis functions in the model. For example, 1 represents the model with 0 basis functions. With 0 basis functions, the model is the mean of the response variable. An alternative model with 9 basis functions has an index of 10.

The tables and graphs for the new model are the same as for the original model. The new tables and graphs are in a new output tab. The same columns that the original analysis stored are also stored for the new model. For example, if the original analysis stored the fits and you select an alternative tree with 9 basis functions, then the new analysis stores the fits in the column Fit[10].

MSPREDICT: Session command to predict from a MARS® model

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

MSPREDICT K

Predicts the response values for new observations with the optimal MARS® model.

K is the identification number of a model. This K is always a positive integer. The first MARS® model that you create in a project is identified as 1, the second is identified as 2, and so on.

Note The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second MARS® model that you made remains model 2.

Specify the values to predict with either CPREDICTORS or KPREDICTORS.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model. If a predictor is categorical, the values in C must match the values in the model. The columns can have any number of values as long as the lengths are equal.

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. If a predictor is categorical, the corresponding value for K must match the values in the model. Use KPREDICTORS from 1 to 3 times with a single instance of the main command.

Argument order of CPREDICTORS and KPREDICTORS

The order of the predictors for the subcommands follows a strict format.

1. Enter continuous predictors first. Use the order of the predictors from the `CONTINUOUS` subcommand from `MARS`.
2. Enter categorical predictors second. Use the order of the predictors from the `CATEGORICAL` subcommand from `MARS`.

For example, suppose you fit a model using a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. The model used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

CPREDICTORS

To use CPREDICTORS, enter the predictors with the same order as the subcommands from the model:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

KPREDICTORS

To use KPREDICTORS, enter specific predictor values. For example, the following table gives values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

To use KPREDICTORS, enter the values using the same order as the subcommands from the model:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Results

TPREDICTION

Displays the prediction table.

Storage

SFIT C
Stores the fit of each row in the column that you specify.

Discover Best Model (Continuous Response)

ATREGRESS: Session command for discovering the best model type for a continuous response

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ATREGRESS
Compares the performance of common types of models from Minitab commands: REGRESS, RTREE, BRTREE, RRFOREST and MARS. All 5 commands model a continuous response with many categorical and continuous predictor variables.

The response variable column must be continuous and must contain at least 2 distinct numeric values.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values in the response variable. For a multiple regression model, Minitab also excludes observations with missing values in a predictor variable.

RESPONSE C
Enter a column that contains 2 or more distinct values of a continuous response. C must be numeric.

CONTINUOUS C...C
Specifies the continuous predictors that are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C
Specifies the categorical predictors that are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Validation options

TEST K
Performs validation with a test set. Minitab Statistical Software randomly selects the rows for the test set. K is the fraction of the data in the test set.

RANDOMIZE K
Specifies the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for the random number generator.

K must be a positive integer. Use 12345 to get the same results as the default setting.

CVTEST K

Performs k-fold cross validation for the multiple regression model, the TreeNet® model, the MARS® model, and the CART® model. The validation of the Random Forests® model uses the out-of-bag statistics.

Minitab Statistical Software randomly selects the data for the folds. K is the number of folds.

SSAMPLE C

With CVTEST:

- The storage column stores the fold identification in C.
- C contains values 1, 2, ..., K.

With TEST:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the 2 values.

Criteria

Use one of the following subcommands to specify the criterion that determines the best model.

- LSE: Use the Least Squared Errors method. This is the default method. Use the HUBER subsubcommand to use the Huber loss function for the TreeNet model.
  - HUBER K: Use the Huber method. K is the switching value from squared loss to absolute loss in the hybrid Huber loss function. K must be from 0.01 to 0.99. K = 0.9 is the default value.
- LAD: Use the Least Absolute Deviation method.

These commands also determine the optimal pruning of the tree for CART® models and the criterion to use to determine the optimal number of trees for the TreeNet® model. If you do not specify any subcommands about the criteria, then the analysis uses the Least Squared Errors method.

Options for multiple regression models

For a multiple regression model, you can specify which terms are in the model and a procedure for term selection. Specification of more than one of the subcommands STEPWISE, FVALIDATION, and NONE stops the analysis.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors can be repeated. Nested terms are not allowed. The model can be nonhierarchical. If you do not specify this subcommand, then the analysis uses all linear terms and all terms of order 2. For examples, go to How to specify the model for GLM on page 1250.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. The term that enters the model has the lowest p-value below 0.15. If no terms can enter the model, the term with the highest p-value greater than 0.15 exits the model. The procedure adds or removes 1 term at a step. The procedure ignores model hierarchy.

FVALIDATION

Uses forward selection to select the terms for the final model. The selection of the final model uses the criterion from the forward selection calculations. At each step, Minitab Statistical Software adds the term with the smallest p-value to the model. The procedure ignores model hierarchy. At the end of each step, Minitab Statistical Software calculates the validation version of the criterion. At the end of the forward selection procedure, the model with the best value of the criterion is the final model.

When you use CVTEST, the procedure repeats forward selection on each fold. The procedure evaluates all the folds at each step and identifies the step with the best overall k-fold value of the criterion. The last part of the procedure is to perform forward selection on the full dataset. The procedure continues until it either fits the full model or until it fits a model that leaves 1 degree of freedom for error. At the end of the forward selection procedure, the model with the best value of the criterion is the final model.

NONE

Uses the model from the TERMS subcommand without any selection procedure. If you do not specify the TERMS subcommand, the NONE subcommand specifies the use of all linear terms and all terms of order 2.

Options for CART® models

Use either MINERROR or STERROR to define how the procedure identifies the optimal pruning of the tree for the CART® model. These two subcommands are mutually exclusive. If you do not specify MINERROR or STERROR, the analysis uses STERROR 1.

MINERROR

Uses the tree with the maximum R^2 or the minimum mean absolute deviation as the optimal tree, depending on the criteria for choosing the best model.

STERROR K

Uses the smallest tree that has a value of the optimality criterion within K standard errors of the best value of the optimality criterion to select the optimal pruning of the tree.

RNOSPLIT K

Do not split a node that is smaller than K. K=10 is the default.

NOTERMINAL K

Do not create a terminal node that is smaller than K. K=3 is the default.

Options for TreeNet® and Random Forests® models

NTREE K K

The first K is the number of trees to build for the TreeNet® model (gradient boosted regression tree). Enter a value from 1 to 5,000. The second K is the number of bootstrap samples and the number of trees to build for the Random Forests® model. Enter a value from 3 to 3,000.

If you do not specify NTREE, Minitab uses 300 for each value of K.

NPRED K K

Specifies the number of predictors randomly selected to evaluate at each node to split. The first K is for TreeNet® models. The second K is for Random Forests® models. K must be an integer between 0 and 100. The following list describes what the subcommand does with different values of K.

- 0: Use the square root of the total number of predictors, rounded down to the nearest integer.
- 100: Use all predictors.
- 0 < K < 100: Use this percentage of the total number of predictors, rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors is 1.

If you do not specify NPRED, then TreeNet® models use all predictors (K = 100) and Random Forests® models use the square root of the number of predictors (K = 0).

BRANDOMIZE K

Specifies the base number to randomly generate subsamples to build a tree. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

Options for TreeNet® models

LRATE K...K

Specifies up to 10 learning rates to evaluate. The learning rate of the boosting procedure protects from overfitting. K must be from 0.0001 to 1.

If you do not specify the LRATE subcommand, then the analysis usually tunes the hyperparameters with 3 values of K: 0.001, 0.1, and max(0.01, 0.1 * min(1.0, N/10000)), where N = number of rows in the response column. If max(0.01, 0.1 * min(1.0, N/10000)) = 0.001 or 0.1, then the analysis tunes the hyperparameters with 0.001, 0.01 and 0.1.

SUBSAMPLE K...K

Specify up to 10 fractions of subsampling. Each iteration selects a different subset of data to construct that tree. Subsampling protects from overfitting. K must be greater than 0 and less than or equal to 1. The default values are 0.5 and 0.7.

For the TreeNet® models, choose either MAXNODE or MAXTDEPTH. If you do not specify either MAXNODE or MAXTDEPTH, then the analysis uses MAXNODE 6.

MAXNODE K (K K)

Specifies up to 3 numbers of maximum terminal nodes per tree. K must be a positive integer between 2 and 2000.

MAXTDEPTH K (K K)

Specifies up to 3 numbers of maximum tree depths to the K^{th} layer. K must be a positive integer between 2 and 1000. MAXTDEPTH tends to give balanced trees.

Options for Random Forests® models

BSIZE K

Sets the bootstrap sample size. You must enter a value of K greater than or equal to 5. If you don't enter the BSIZE subcommand or if you enter a value that is greater than the training data size, then Minitab uses a sample size equal to the training data size.

NOSPLIT K (K K)

Does not split a node that is smaller than K. Specifies up to 3 integers greater than or equal to 2. Each integer produces a distinct Random Forests® model. If you do not specify the NOSPLIT subcommand, then the analysis uses K = 5. If you specify a K larger than the maximum value for the data, then the analysis uses the maximum value.

Options for MARS® models

BFUNCTIONS K

Specifes the number of basis functions to evaluate, $1 \leq K \leq 200$. If you do not specify this subcommand, then the analysis uses K = 30.

MINOBS K

Specifies to use an algorithm to determine the minumum number of data points between changes in the basis functions or specifies the minumum number.

K = 0

The analysis uses sample size and model complexity to automatically select a value.

1 ≤ K ≤ N

A value of 1 indicates that consecutive data points are eligible to be points where the basis function changes. The value of 1 allows the most rapid changes in the model predictions. Use larger values to create smoother, potentially less locally-accurate models to explore more general relationships. N is the number of rows in the column for the response variable. Values larger than the sample size result in an error.

Use one of the following subcommands to set the predictor interaction options. If no subcommand from this set is issued, NINTERACT is the default. To evaluate a MARS® model with higher-order interactions, fit the model with the MARS command.

NINTERACT

Do not allow predictor interactions when the analysis evaluates the basis functions. Minitab uses only the additive model.

SINTERACT 2

Allow second-order interactions when the analysis evaluates the basis functions.

Storage

SMODEL

Stores the model object to use in subsequent analyses. If you do not store the model, then you cannot perform analyses like prediction, selection of an alternative model, or the addition of partial dependence plots.

SFIT C

Stores the fitted response for each record in C.

SRESID C

Stores the residual for each record in C.

ARTMODEL: Session command for selecting an alternative model from **Discover Best Model (Continuous Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ARTMODEL K

Produces results for alternative models after ATREGRESS identifies the best model.

K is a positive integer that represents a **Discover Best Model (Continuous Response)** object ID. The first model from **Discover Best Model (Continuous Response)** that you create in a project has the object ID 1. The second model from **Discover Best Model (Continuous Response)** that you create with the command has the object ID 2, and so on.

The tables and graphs for the new model are displayed in a new output tab. The same columns that the original analysis stored are also stored for the new model. The storage for the new model is in a new column. For example,

if the original analysis stored the fitted values in a column titled "Fit," then the new analysis names an empty column "Fit_1" and stores the fitted values in that column.

The type of object ID that ARTMODEL creates is the same as the type of model in the results. The type of object ID determines which command to use to predict from a model. For TreeNet® models, you can also add partial dependence plots. Use the following commands to do further analysis on each type of model.

Type	Analysis	Command
Multiple regression	Predict the response for new data	MRPREDICT
CART®	Predict the response for new data	RTPREDICT
Random Forests®	Predict the response for new data	RRPREDICT
TreeNet®	Predict the response for new data	BRPREDICT
MARS®	Predict the response for new data	MSPREDICT

For purposes of using a command such as MRPREDICT or BRPREDICT, the object ID that comes from the ARTMODEL command is the next object ID of that type. For example, if ARTMODEL creates the first CART® regression model in the Minitab project, then the first CART® regression model has the object ID 1. If you then use RTREE to create a different CART® regression model, then that model has the object ID 2.

Results for an existing model

To produce the results for the multiple regression model, do not enter any subcommands. Use the appropriate subcommand to produce the results for a TreeNet® model, a Random Forests® model, a MARS® model, or a CART® model.

TN K

Produces results for the TreeNet® model that K specifies. K is the value in the Model column of the Optimization of Hyperparameters table that shows the different TreeNet® models. If this table is not in the results, then the ID is 1.

RF K

Produces results for the Random Forests® model that K specifies. K is the value in the Model column of the Optimization of Hyperparameters table that shows the different Random Forests® models. If this table is not in the results, then the ID is 1.

MARS K

Produces results for the MARS® model that K specifies. K values are 1 more than the number of basis functions in the model. For example, 1 represents the model with 0 basis functions. With 0 basis functions, the model is the mean of the response variable. An alternative model with 9 basis functions has an index of 10.

CART

Produces results for the best CART® model from the search for the best type of model.

Results for a new model

To produce results for a TreeNet® model, a Random Forests® model, or a MARS® model, specify which type of model to create and then specify all the hyperparameters for the model. For TreeNet® models and Random Forests® models, you can enter multiple values of the hyperparameters. If you enter multiple values for more than 1 hyperparameter, then the procedure evaluates all of the possible combinations of the hyperparameters and produces full results for the best model.

TN

Produces results for the TreeNet® model that you specify with the subsubcommands.

LRATE K...K

Specifies up to 10 learning rates to evaluate. You can specify values from 0.0001 to 1.

SUBSAMPLE K...K

Specifies up to 10 subsample proportions to evaluate. You can specify positive values that are less than or equal to 1.

MAXNODE K (K K)

Specifies up to 3 maximum numbers of terminal nodes to evaluate. You can specify values from 2 to 2,000. You cannot use MAXNODE and MAXTDEPTH together.

MAXTDEPTH K (K K)

Specifies up to 3 maximum tree depths to evaluate. You can specify values from 2 to 1,000. You cannot use MAXNODE and MAXTDEPTH together.

NTREE K

Specifies the maximum number of trees to build. Specify a single value from 1 to 5,000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

RF

Produces results for the Random Forests® model that you specify with the subsubcommands. If you specify more than one argument for NPRED or NOSPLIT, the procedure evaluates all of the possible combinations of the hyperparameters and produces full results for the best model.

NTREE K

Specifies K bootstrap samples to grow K trees. You must enter a value between 3 and 3000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

NOSPLIT K (K K)

Do not split a node if the node size is less than K.

BSIZE K

Specifies the bootstrap sample size. You must enter a value greater than or equal to 5. If you don't enter the BSIZE subsubcommand or if you enter a value that is greater than the training data size, Minitab uses a sample size equal to the training data size. This subsubcommand is optional.

MARS

Produces results for the MARS® model that you specify with the subsubcommands. Use NINTERACT or SINTERACT 2 to specify the predictor interaction options. To explore higher order interactions or to control interactions more discretely, use the MARS command to fit a new model.

BFUNCTIONS K

Set the maximum number of basis functions to evaluate in the construction of the model. K must be a positive integer of 200 or less. K = 30 is the default value.

MINOBS K

Specifies to use an algorithm to determine the minumum number of data points between changes in the basis functions or specifies the minumum number.

K = 0

The analysis uses sample size and model complexity to automatically select a value.

1 ≤ K ≤ N

A value of 1 indicates that consecutive data points are eligible to be points where the basis function changes. The value of 1 allows the most rapid changes in the model predictions. Use larger values to create smoother, potentially less locally-accurate models to explore more general relationships. N is the number of rows in the column for the response variable. Values larger than the sample size result in an error.

NINTERACT

Do not allow predictor interactions in the model. Minitab uses only the additive model.

SINTERACT 2

Allow predictor interactions up to the 2nd order when building the model.

ARTPREDICT: Session command to predict from a model in the results of **Discover Best Model (Continuous Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ARTPREDICT K

Predicts the response values for new observations with the model that **Discover Best Model (Continuous Response)** saves.

K is the identification number of a model that the ATREGRESS command saves. This K is always a positive integer. The first model from the ATREGRESS command that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model.

If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on whether the model uses all of the predictors from the original analysis. CART®, MARS®, Random Forests®, and TreeNet® models always use all of the predictors from the original subcommands for the analysis. A multiple regression model that uses a selection procedure can have fewer predictors than in the original subcommands for the analysis.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the results contain a multiple regression model that eliminates some predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATREGRESS command, the predictor selection procedure returns a model without C3 and C4 as predictors. Then, you can run the prediction command with the CPREDICTORS subcommand.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the tree.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on whether the model uses all of the predictors from the original analysis. CART®, MARS®, Random Forests®, and TreeNet® models always use all of the predictors from the original subcommands for the analysis. A multiple regression model that uses a selection procedure can have fewer predictors than in the original subcommands for the analysis.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from ATREGRESS.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from ATREGRESS.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1

Predictor	Value
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the results contain a multiple regression model that eliminates some predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from ATREGRESS.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from ATREGRESS.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATREGRESS command, the predictor selection procedure returns a model without C3 and C4 as predictors. Then, you can run the prediction command with the KPREDICTORS subcommand.

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Storage and results for all types of models

The table is available when you use CPREDICTORS or KPREDICTORS. Storage is available only when you use CPREDICTORS.

TPREDICTION
Displays the prediction table.

SFIT C
Stores the predicted responses in C.

Additional storage and results for multiple regression models

The table is available when you use CPREDICTORS or KPREDICTORS. Storage is available only when you use CPREDICTORS.

TEQUATION
Displays the model equation. The results provide a separate equation for each level of each categorical predictor. If you have only categorical variables, the results show a single equation.

SSEFIT C
Stores the standard error of the predicted responses in C.

SCLIMITS C C
Stores the lower and upper confidence limits of the predicted responses in C and C.

SPLIMITS C C
Stores the lower and upper prediction limits of the predicted responses in C and C.

Additional storage for CART® models

Storage is available only when you use CPREDICTORS.

STNID C
Stores the terminal node (cluster) ID for each new observation.

MRPREDICT: Session command to predict from a multiple regression model that was an alternative model from **Discover Best Model (Continuous Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

MRPREDICT K
Predicts response values for new observations from the multiple regression model object ID that K identifies. The ARTMODEL command with no subcommands creates a multiple regression model object.

This K is always a positive integer. The first multiple regression model object that you create in a project is identified as 1, the second object is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model object, the second model object that you made remains model object 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model. If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on if a selection algorithm removes any predictors from the model. Predictor selection does not occur when the ATREGRESS command includes the NONE subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the model does not include all of the predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATREGRESS command, the predictor selection algorithm gave a model without C3 and C4 as predictors. Then, you can run the prediction command with the CPREDICTORS subcommand.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. KPREDICTORS can be repeated up to three times.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on if a selection algorithm removes any predictors from the model. Predictor selection does not occur when the ATREGRESS command includes the NONE subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the model does not include all of the predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATREGRESS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATREGRESS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATREGRESS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATREGRESS command, the predictor selection algorithm gave a model without C3 and C4 as predictors. Then, you can run the prediction command with the KPREDICTORS subcommand.

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Results

Displays a variety of tables in the output.

TEQUATION
Displays the model equation. The results provide a separate equation for each level of each categorical predictor. If you have only categorical variables, the results show a single equation.

TPREDICTION
Displays the prediction table.

Storage

Stores information related to the predicted responses when you use the CPREDICT subcommand.

SFIT C
Stores the predicted responses in C.

SSEFIT C
Stores the standard error of the predicted responses in C.

SCLIMITS C C
Stores the lower and upper confidence limits of the predicted responses in C and C.

SPLIMITS C C
Stores the lower and upper prediction limits of the predicted responses in C and C.

Discover Best Model (Binary Response)

ATCLASS: Session command for discovering the best model type for a binary response

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ATCLASS
Compares the performance of 4 common types of models from Minitab commands: GZLM, CTREE, BCTREE, and RCFOREST. The models all use a binary response with many categorical and continuous predictor variables.

The response variable column must contain two distinct values that are text or numbers.

The predictors can be continuous predictors or categorical predictors. Continuous predictors must have numeric values. Categorical predictors can have numeric or text values. You must use at least one of the predictor subcommands, either CONTINUOUS or CATEGORICAL.

All columns in the analysis must have the same number of rows. Minitab excludes all observations with missing values in the response variable. For a binary logistic regression model, Minitab also excludes observations with missing values in a predictor variable.

RESPONSE C

Enter a column that contains 2 distinct values of a categorical response. C can be text or numeric.

EVENT K

Specifies the event value for a binary response. K must be 1 of the 2 response values. Enclose a text value of K in quotation marks.

CONTINUOUS C...C

Specifies the continuous predictors that are in columns C...C. The columns must be numeric and must have the same number of rows as the response column.

CATEGORICAL C...C

Specifies the categorical predictors that are in columns C...C. The columns can be numeric or text and must have the same number of rows as the response column.

Validation options

TEST K

Performs validation with a test set. Minitab Statistical Software randomly selects the rows for the test set. K is the fraction of the data in the test set.

RANDOMIZE K

Specifies the base number for randomly splitting the entire data set into training and test sets. K sets the starting point for the random number generator.

K must be a positive integer. Use 12345 to get the same results as the default setting.

CVTEST K

Performs k-fold cross validation for the binary logistic regression model, the TreeNet® model, and the CART® model. The validation of the Random Forests® model uses the out-of-bag statistics.

Minitab Statistical Software randomly selects the data for the folds. K is the number of folds.

SSAMPLE C

With CVTEST:

- The storage column stores the fold identification in C.
- C contains values 1, 2, ..., K.

With TEST:

- The storage column, Sample_Id, stores the training and test sample assignments in C.
- C is a binary text column that contains Training and Test as the 2 values.

Criteria

Use one of the following subcommands to specify the criterion that determines the best model.

- MAXLKH: Use the maximum likelihood.
- ROC: Use the maximum area under the ROC curve.
- MISCLASS: Use the minimum misclassification rate.

These commands also determine the optimal pruning of the tree for CART® models and the criterion to use to determine the optimal number of trees for the TreeNet® model. If you do not specify any subcommands about the criteria, then the analysis uses the Maximum Likelihood method.

Options for binary logistic regression models

For a binary logistic regression model, you can specify which terms are in the model and a procedure for term selection. Specification of more than one of the subcommands STEPWISE, FVALIDATION, and NONE stops the analysis.

TERMS *termlist*

Specifies the model terms. Terms must be legal cross-terms. Only continuous predictors can be repeated. Nested terms are not allowed. The model can be nonhierarchical. If you do not specify this subcommand, then the analysis uses all linear terms and all terms of order 2. For examples, go to How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC on page 1248.

STEPWISE

Specifies a stepwise model selection procedure that uses both forward selection and backward elimination. The term that enters the model has the lowest p-value below 0.15. If no terms can enter the model, the term with the highest p-value greater than 0.15 exits the model. The procedure adds or removes 1 term at a step. The procedure ignores model hierarchy.

FVALIDATION

Uses forward selection to select the terms for the final model. The selection of the final model uses the criterion from the forward selection calculations. When you use FVALIDATION, you must also use TEST. At each step, Minitab Statistical Software adds the term with the smallest p-value to the model. The procedure ignores model hierarchy. At the end of each step, Minitab Statistical Software calculates the validation version of the criterion. At the end of the forward selection procedure, the model with the best value of the criterion is the final model.

The procedure continues until it either fits the full model or until it fits a model that leaves 1 degree of freedom for error.

NONE

Uses the model from the TERMS subcommand without any selection procedure. If you do not specify the TERMS subcommand, the NONE subcommand specifies the use of all linear terms and all terms of order 2.

Options for binary logistic regression models and *TreeNet*® models

CTHRESHOLD K

For a binary response, K specifies the threshold value such that if the probability of the event > K, then assign the event level as the predicted response. Otherwise, assign the non-event level as the predicted response.

K must be between 0 and 1; K = 0.5 is the default. K = 0 indicates that the threshold equals the estimated event probability of the data when no model is built.

Options for CART® models

Use one of the following subcommands to specify the node splitting method.

- GINI: Use the Gini method. This is the default method.
- ENTROPY: Use the Entropy method.

The following subcommands are available to select the optimal pruning of the tree.

- MINCOST: Use minimum cost to select the optimal pruning of the tree.
- STERROR K : Use within K standard errors of the minimum cost to select the optimal pruning of the tree.

If you do not specify the node splitting method, then the analysis uses GINI and STERROR 1.

The following subcommands are available to control tree growth due to sample size or node complexity.

- CNOSPLIT K: Do not split a node that is smaller than K. K=10 is the default.
- NOTERMINAL K: Do not create a terminal node that is smaller than K. K=3 is the default.

Options for TreeNet® and Random Forests® models

NTREE K K

The first K is the number of trees to build for the TreeNet® model (gradient boosted regression tree). Enter a value from 1 to 5,000. The second K is the number of bootstrap samples and the number of trees to build for the Random Forests® model. Enter a value from 3 to 3,000.

If you do not specify NTREE, Minitab uses 300 for each value of K.

NPRED K K

Specifies the number of predictors randomly selected to evaluate at each node to split. The first K is for TreeNet® models. The second K is for Random Forests® models. K must be an integer between 0 and 100. The following list describes what the subcommand does with different values of K.

- 0: Use the square root of the total number of predictors, rounded down to the nearest integer.
- 100: Use all predictors.
- 0 < K < 100: Use this percentage of the total number of predictors, rounded down to the nearest integer. If (total number of predictors) x (K / 100) < 1, the number of predictors is 1.

If you do not specify NPRED, then TreeNet® models use all predictors (K = 100) and Random Forests® models use the square root of the number of predictors (K = 0).

BRANDOMIZE K

Specifies the base number to randomly generate subsamples to build a tree. K must be a positive integer. K also controls every randomly selected subset of the m total predictors to generate the splitting rule for an internal node, if applicable.

Options for TreeNet® models

LRATE K...K

Specifies up to 10 learning rates to evaluate. The learning rate of the boosting procedure protects from overfitting. K must be from 0.0001 to 1.

If you do not specify the LRATE subcommand, then the analysis usually tunes the hyperparameters with 3 values of K: 0.001, 0.1, and max(0.01, 0.1 * min(1.0, N/10000)), where N = number of rows in the response column. If max(0.01, 0.1 * min(1.0, N/10000)) = 0.001 or 0.1, then the analysis tunes the hyperparameters with 0.001, 0.01 and 0.1.

SUBSAMPLE K...K

Specify up to 10 fractions of subsampling. At each iteration, the procedure selects a different subset that contains this fraction of the data to construct a tree. Subsampling protects from overfitting. K must be greater than 0 and less than or equal to 1. The default values are 0.5 and 0.7.

For the TreeNet® models, choose either MAXNODE or MAXTDEPTH. If you do not specify either MAXNODE or MAXTDEPTH, then the analysis uses MAXNODE 6.

MAXNODE K (K K)
Specifies up to 3 numbers of maximum terminal nodes per tree. K must be a positive integer between 2 and 2000.

MAXTDEPTH K (K K)
Specifies up to 3 numbers of maximum tree depths to the K^{th} layer. K must be a positive integer between 2 and 1000. MAXTDEPTH tends to give balanced trees.

Options for Random Forests® models

BSIZE K
Sets the bootstrap sample size. You must enter a value of K greater than or equal to 5. If you don't enter the BSIZE subcommand or if you enter a value that is greater than the training data size, then Minitab uses a sample size equal to the training data size.

NOSPLIT K (K K)
Does not split a node that is smaller than K. Specifies up to 3 integers greater than or equal to 2. Each integer produces a distinct Random Forests® model. If you do not specify the NOSPLIT subcommand, then the analysis uses K = 2. If you specify a K larger than the maximum value for the data, then the analysis uses the maximum value.

Storage

SMODEL
Stores the model object to use in subsequent analyses. You can predict, select an alternative model, or add more partial dependence plots only if you store the model object.

SFIT C
Stores the fitted event probability for each record in C. These fitted values are for TreeNet® models and binary logistic regression models. These fitted values are not for CART® models or Random Forests® models.

SPCLASS C
Stores the predicted class for each record in C.

SEPROB C C
Stores the probability of the event and the probability of the nonevent.

ACTMODEL: Session command for selecting an alternative model from **Discover Best Model (Binary Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ACTMODEL K
After the analysis identifies the best model with ATCLASS, ACTMODEL produces results for alternative models.

- With none of the subcommands that specify a model type, the analysis produces results for the best binary logistic model from the search.
- With the CART subcommand, the analysis produces results for the best CART® model from the search for the best model.
- With the TN subcommand and an object ID that identifies the model, the analysis produces results for the best TreeNet® model from the search for the best model.

- With the RF subcommand and an object ID that identifies the model, the analysis produces results for the best Random Forests® model from the search for the best model.
- With the TN subcommand or the RF subcommand in addition to the subsubcommands that specify details about the model type, ACTMODEL explores additional models.

K is a positive integer that represents a **Discover Best Model (Binary Response)** object ID. The first **Discover Best Model (Binary Response)** that you create in a project has the object ID 1. The second **Discover Best Model (Binary Response)** that you create with the command has the object ID 2, and so on.

The tables and graphs for the new model are displayed in a new output tab. The same columns that the original analysis stored are also stored for the new model. The storage for the new model is in a new column. For example, if the original analysis stored the fitted values in a column titled "Fit," then the new analysis names an empty column "Fit_1" and stores the fitted values in that column.

The type of object ID that ACTMODEL creates is the same as the type of model in the results. The type of object ID determines which command to use to predict from a model. For TreeNet® models, you can also add partial dependence plots. Use the following commands to do further analysis on each type of model.

Type	Analysis	Command
Binary logistic regression	Predict the response for new data	LRPREDICT
CART®	Predict the response for new data	CTPREDICT
Random Forests®	Predict the response for new data	RCPREDICT
TreeNet®	Predict the response for new data	BCPREDICT
	Add one-predictor partial dependence plots	BCGOPART
	Add two-predictor partial dependence plots	BCGTPART

For purposes of using a command such as LRPREDICT or BCPREDICT, the object ID that comes from the ACTMODEL command is the next object ID of that type. For example, if ACTMODEL creates the first CART® classification model in the Minitab project, then that CART® Classification model has the object ID 1. If you then use CTREE to create a different CART® classification model, then that model has the object ID 2.

Results for an existing model

To produce the results for the binary logistic regression model, do not enter any subcommands. Use the appropriate subcommand to produce the results for a TreeNet® model, a Random Forests® model, or a CART® model.

TN K

Produces results for the TreeNet® model that K specifies. K is the value in the Model column of the Optimization of Hyperparameters table that shows the different TreeNet® models. If this table is not in the results, then the ID is 1.

RF K

Produces results for the Random Forests® model that K specifies. K is the value in the Model column of the Optimization of Hyperparameters table that shows the different Random Forests® models. If this table is not in the results, then the ID is 1.

CART

Produces results for the best CART® model from the search for the best type of model.

Results for a new model

To produce results for a TreeNet® model or a Random Forests® model, specify which type of model to create and then specify all the hyperparameters for the model. If you enter multiple values for more than 1 hyperparameter, then the procedure evaluates all of the possible combinations of the hyperparameters and produces full results for the best model.

TN

Produces results for the TreeNet® model that you specify with the subsubcommands.

LRATE K...K

Specifies up to 10 learning rates to evaluate. You can specify values from 0.0001 to 1.

SUBSAMPLE K...K

Specifies up to 10 subsample proportions to evaluate. You can specify positive values that are less than or equal to 1.

MAXNODE K (K K)

Specifies up to 3 maximum numbers of terminal nodes to evaluate. You can specify values from 2 to 2,000. You cannot use MAXNODE and MAXTDEPTH together.

MAXTDEPTH K (K K)

Specifies up to 3 maximum tree depths to evaluate. You can specify values from 2 to 1,000. You cannot use MAXNODE and MAXTDEPTH together.

NTREE K

Specifies the maximum number of trees to build. Specify a single value from 1 to 5,000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

RF

Produces results for the Random Forests® model that you specify with the subsubcommands. If you specify more than one argument for NPRED or NOSPLIT, the procedure evaluates all of the possible combinations of the hyperparameters and produces full results for the best model.

NTREE K

Specifies K bootstrap samples to grow K trees. You must enter a value between 3 and 3000.

NPRED K (K K)

Specifies the number of predictors randomly selected to evaluate at each node to split. K must be a positive integer. If K is greater than the number of predictors, then K counts as being equal to the number of predictors.

NOSPLIT K (K K)

Do not split a node if the node size is less than K.

BSIZE K

Specifies the bootstrap sample size. You must enter a value greater than or equal to 5. If you don't enter the BSIZE subsubcommand or if you enter a value that is greater than the training data size, Minitab uses a sample size equal to the training data size. This subsubcommand is optional.

ACTPREDICT: Session command to predict from a model in the results of **Discover Best Model (Binary Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

ACTPREDICT K

Predicts the response values for new observations with the model that **Discover Best Model (Binary Response)** saves.

K is the identification number of a model that the ATCLASS command saves. This K is always a positive integer. The first model from the ATCLASS command that you create in a project is identified as 1, the second is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model, the second model that you made remains model 2.

CPREDICTORS C...C

Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model.

If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on whether the model uses all of the predictors from the original analysis. CART®, Random Forests®, and TreeNet® models always use all of the predictors from the original subcommands for the analysis. A binary logistic regression model that uses a selection procedure can have fewer predictors than in the original subcommands for the analysis.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATCLASS command.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the results contain a binary logistic regression model that eliminates some predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATCLASS command.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATCLASS command, the predictor selection procedure gave a model without C3 and C4 as predictors. Then, you can run the prediction command with the CPREDICTORS subcommand.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the tree.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on whether the model uses all of the predictors from the original analysis. CART®, Random Forests®, and TreeNet® models always use all of the predictors from the original subcommands for the analysis. A binary logistic regression model that uses a selection procedure can have fewer predictors than in the original subcommands for the analysis.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from ATCLASS.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from ATCLASS.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the results contain a binary logistic regression model that eliminates some predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from ATCLASS.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from ATCLASS.

For example, suppose you have a worksheet with predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATCLASS command, the predictor elimination algorithm gave a model without C3 and C4. Then, you can run the prediction command with the KPREDICTORS subcommand.

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Storage and results for all types of models

The table is available when you use CPREDICTORS or KPREDICTORS. Storage is available only when you use CPREDICTORS.

TPREDICTION
Displays the prediction table.

SEPROB C C
Stores the predicted event and non-event probabilities in C and C.

SPCLASS C
Stores the predicted class in C.

Additional storage and results for binary logistic regression models

The table is available when you use CPREDICTORS or KPREDICTORS. Storage is available only when you use CPREDICTORS.

TEQUATION
Displays the model equation. The results provide a separate equation for each level of each categorical predictor. If you have only categorical variables, the results show a single equation.

SFIT C
Stores the predicted event probability in C.

Additional storage for TreeNet® models

Storage is available only when you use CPREDICTORS.

SFIT C
Stores the predicted event probability in C.

Additional storage for CART® models

Storage is available only when you use CPREDICTORS.

STNID C
Stores the terminal node (cluster) ID for each new observation.

LRPREDICT: Session command to predict from a binary logistic regression model with **Discover Best Model (Binary Response)**

Note This command is available with the **Predictive Analytics Module**. Click here for more information about how to activate the module.

LRPREDICT K
Predicts response values for new observations from the binary logistic regression model object ID that K identifies. The ACTMODEL command with no subcommands creates a binary logistic regression model object.

This K is always a positive integer. The first binary logistic regression model object that you create in a project is identified as 1, the second object is identified as 2, and so on. The number for a model remains the same even if you delete the output for another model from the Navigator. For example, if you delete the output for the first model object, the second model object that you made remains model object 2.

CPREDICTORS C...C
Specifies the predictor values in columns C...C. The number of columns must equal the number of predictors used in the model. If a predictor is categorical, the values in C must match the values in the model. Which predictors to enter depends on if a selection algorithm removes any predictors from the model. Predictor selection does not occur when the ATCLASS command includes the NONE subcommand.

All predictors from the original subcommands
The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATCLASS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

CPREDICTORS C6 C2 C3 C4 C5 C7 C8

Fewer predictors than in the original subcommands

If the model does not include all of the predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATREGRESS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The ATCLASS command uses the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATCLASS command, the predictor selection algorithm gave a model without C3 and C4 as predictors. Then, you can run the prediction command with the CPREDICTORS subcommand.

Enter the predictors in the following order:

CPREDICTORS C6 C2 C5 C7 C8

KPREDICTORS K...K

Specifies the predictor values in K...K. The number of constants must equal the number of predictors used in the model. KPREDICTORS can be repeated up to three times.

If a predictor is categorical, the corresponding value for K must match the values in the model. Which predictors to enter depends on if a selection algorithm removes any predictors from the model. Predictor selection does not occur when the ATCLASS command includes the NONE subcommand.

All predictors from the original subcommands

The order of the predictors follows a strict sequence:

1. Enter continuous predictors. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors. Use the order of the predictors from the CATEGORICAL subcommand from the ATCLASS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A

Predictor	Value
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATCLASS command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 100 'A' 'Blue' 'Urgent' 'New'

Fewer predictors than in the original subcommands

If the model does not include all of the predictors, then omit the eliminated predictors.

1. Enter continuous predictors that are in the model. Use the order of the predictors from the CONTINUOUS subcommand from the ATCLASS command.
2. Enter categorical predictors that are in the model. Use the order of the predictors from the CATEGORICAL subcommand from the ATCLASS command.

For example, suppose the worksheet has predictors in C2-C8. Predictors C2, C3, and C6 are continuous and predictors C4, C5, C7, and C8 are categorical. You use the ATCLASS command to create a model. The following table gives the values to predict:

Predictor	Value
C2	12.1
C3	100
C4	A
C5	Blue
C6	-45.3
C7	Urgent
C8	New

The ATCLASS command used the following subcommands:

- CONTINUOUS C6 C2 C3
- CATEGORICAL C4 C5 C7 C8

When you run the ATCLASS command, the predictor selection algorithm gave a model without C3 and C4 as predictors. Then, you can run the prediction command with the KPREDICTORS subcommand.

Enter the predictors in the following order:

KPREDICTORS -45.3 12.1 'Blue' 'Urgent' 'New'

Results

Displays a variety of tables in the output.

TEQUATION
Displays the model equation. The results provide a separate equation for each level of each categorical predictor. If you have only categorical variables, the results show a single equation.

TPREDICTION
Displays the prediction table.

Storage

Stores information related to the predicted responses when you use the CPREDICT subcommand.

SEPROB C C
Stores the predicted event and non-event probabilities in C and C.

SPCLASS C
Stores the predicted class in C.

SFIT C
Stores the predicted event probability in C.

Multivariate Analysis

PCA: Session command for performing principal components analysis

PCA C...C

Performs principal component analysis (PCA).

PCA is a data reduction technique that is used to identify a small set of variables that account for a large proportion of the total variance in the original variables. Components can be calculated from the correlation matrix (the default) or the covariance matrix. Output consists of the eigenvalues (that is, the variances of the principal components), the proportion and cumulative proportion of the total variance explained by each principal component, and the coefficients for each principal component.

COVARIANCE

Calculates the principal components using the covariance matrix. The default is the correlation matrix.

NCOMP K

If there are p variables on the PCA line, then the default is to calculate all p principal components. But if p is large, you may want just the first few. NCOMP says how many components to calculate.

Even if you use NCOMP, Minitab still prints all p eigenvalues.

COEFFICIENTS C...C

Stores the coefficients of the principal components in the columns. The number of columns given on COEFFICIENTS must be less than or equal to the number of principal components calculated.

SCORES C...C

Stores the principal components scores in the columns. The number of columns on SCORES must be less than or equal to the number of principal components calculated.

EIGEN C

Stores the eigenvalues in one column. The eigenvalues are stored in numerical order from largest to smallest.

DISTANCE C

Stores the Mahalanobis distance for the data points in one column.

GSCREE

Displays a Scree plot (also called eigenvalue profile plot). Minitab plots the eigenvalue associated with a principal component versus the number of the component.

GSCORE

Plots the scores for the second principal component (y-axis) versus the scores for the first principal component (x-axis). To create plots for other components, store the scores with the SCORES subcommand and use the PLOT on page 1020 command.

GLOADING

Plots of the loadings for the second principal component (y-axis) versus the loadings for the first principal component (x-axis). A line is drawn from each loading to the (0, 0) point.

GBIPLOT

Displays a biplot, a combined score and loadings plot, for the first two components.

GOUTLIER

Plots the Mahalanobis distance for each data point. This distance is the difference between the multivariate space's centroid and the data point. Use this plot to identify outliers in a multivariate space because it takes into account the different scales between variables and the correlations among them.

FACTOR: Session command for performing a factor analysis

FACTOR [C...C]

Performs factor analysis. Initial factor extraction may be done using principal components (default) or maximum likelihood. Input can be the data, a correlation or covariance matrix, or the loadings from a previous analysis.

You can rotate the initial solution using varimax, quartimax, equamax, or orthomax.

Displayed output consists of the loadings (factor pattern), communalities, and the coefficients for the factor scores. You can store the loadings, factor score coefficients, factor scores, and residual matrix.

Extraction

NFACTORS K

Extracts K factors.

NFAC is required if you specify ML for the initial extraction. If you do not use NFACTORS with principal components, Minitab sets NFACTORS equal to the number of variables in the data set.

If, with ML, you specify a value of K that is too large, the solution will not be unique. Minitab will print a message in this case. Specifically, if there are p variables on FACTOR, then there are p(p - 1)/2 parameters in the correlation matrix and pK parameters in the factor model fit by ML. If K > (p – 1)/2 then the number of parameters in the factor model is greater than the number in the correlation matrix. The solution is not unique. Rotation will not help in this case.

ML

Uses maximum likelihood for the initial solution. The default is to use principal components. If you use ML, you must specify the number of factors you want with NFACTORS.

Rotation

FACTOR provides four orthogonal rotations. VMAX specifies a varimax rotation, QMAX a quartimax rotation, EMAX an equamax rotation, and OMAX an orthomax rotation with gamma = K.

The first three rotations are actually special cases of OMAX with the appropriate choice for gamma; gamma = 1 is varimax, gamma = 0 is quartimax, and gamma = NFACTORS/2 is equamax.

VMAX

Specifies varimax rotation.

QMAX
Specifies quartimax rotation.

EMAX
Specifies equamax rotation.

OMAX K
Specifies orthomax rotation with K gamma.

Options

CORRELATION ***[M]***
Factors the correlation matrix. This is the default. Use COVARIANCE to factor the covariance matrix.

If you specify a stored correlation matrix on CORRELATION, Minitab calculates the loadings and coefficients from the stored matrix. In this case, any data specified on FACTOR is ignored, and thus scores can not be calculated.

COVARIANCE ***[M]***
Factors the covariance matrix. The default is to factor the correlation matrix. See CORRELATION.

If you specify a stored covariance matrix on COVARIANCE, Minitab calculates the loadings and coefficients from the stored matrix. In this case, any data specified on FACTOR is ignored, and thus scores can not be calculated.

COVARIANCE cannot be used with ML.

PRIORS C
FACTOR uses these in the algorithm for ML, as the initial values for the communalities. C should contain one value for each variable.

The default is to use the squared multiple correlations of each variable with all the other variables. Suppose R is the correlation matrix of the variables. Suppose rii is the i^{th} diagonal element of R**–1 then initial value of the i^{th} communality is (1-1/rii).

MAXITERATIONS K
Specifies the maximum number of iterations for ML. The default is 20.

CONVERGENCE K
Specifies the criterion for convergence in ML. The ML algorithm converges when the uniquenesses do not change very much. The value of K is the size of the smallest change. By default, K is 0.005.

INLOADINGS C...C
Uses specified columns as loadings.

INLOADINGS allows you to try different rotations without having to recalculate the initial solution each time. It is particularly useful if you use ML for the initial solution, since this can take a long time to calculate. If you want the rotated loadings and rotation matrix only, you do not have to specify the data on FACTOR. If you want the coefficients, factor scores, or residual matrix, then you must specify the data.

You can also use INLOADINGS to predict factor scores for new data from previously calculated loadings, stored with the subcommand LOADINGS. Specify the new data on FACTOR and the calculated loadings on INLOADINGS.

Graphs

GSCREE

Displays a Scree plot, also called an eigenvalue profile plot. Minitab plots the eigenvalue associated with a factor versus the number of the factor.

GSCORE

Plots the scores for the second factor (y-axis) versus the scores for the first factor (x-axis). To create plots for other factors, store the scores with the SCORES subcommand and use the PLOT on page 1020 command. (Note: Scores must be calculated from raw data. Therefore, you must not specify a matrix with either the CORRELATION or COVARIANCE subcommands if you want to generate this graph.)

GLOADING

Plots of the loadings for the second factor (y-axis) versus the loadings for the first factor (x-axis). A line is drawn from each loading to the (0, 0) point.

To create simple loading plots for other factors, store the loadings with the LOADINGS subcommand and use the PLOT on page 1020 command. If you want to connect the loading point to the zero point, add a zero to the bottom of each column of loadings in the worksheet, then add lines connecting the loading points to the zero point with the graph editor.

GBIPLOT

Displays a biplot, a combined score and loadings plot, for the first two factors.

Storage

EIGEN C *M*

Stores the eigenvalues and eigenvectors of the matrix that was factored, either the correlation or covariance matrix.

EIGEN can be used only when the initial extraction is done using principal components. The eigenvalues are stored in C, in numerical order starting with the largest, and the eigenvectors in M, one vector per column of **M**, in the same order as the eigenvalues. You can then PLOT the eigenvalues to see a scree plot.

LOADINGS C...C

Stores the loadings in C...C. There must be one C for each factor. These can later be input with the subcommand INLOADINGS. They can also be plotted with PLOT. If a rotation was specified, the loadings for the rotated solution are stored.

COEFFICIENTS C...C

Stores the factor score coefficients in C...C. There must be one C for each factor. These can then be plotted with PLOT. If a rotation was specified, the coefficients for the rotated solution are stored.

SCORES C...C

Stores the factor scores in C...C. There must be one C for each factor. These can then be plotted with PLOT. If a rotation was specified, the scores for the rotated solution are stored. To use SCORES, you must specify data as input on the FACTOR command line, and you must not specify a matrix with either the CORRELATION or COVARIANCE subcommands.

Minitab calculates the factor scores from the original variables and the factor score coefficients. First the variables are centered by subtracting means. Then to get the scores corresponding to factor i, the centered variables are multiplied by the factor score coefficients corresponding to factor i. These are some times called standard scores.

RESIDUAL *M*

Stores the residual matrix. If the matrix, A, was factored, the residual matrix is (A – LL'), where L is the matrix of loadings (after rotation if a rotation was done). Note that the residual matrix for the initial and rotated solutions are the same.

ROTMAT *M*

Stores the matrix used to rotate the initial loadings. If **L** is the matrix of initial loadings, then **LM** is the matrix of rotated loadings.

Results

SORT K

Displays the loadings in sorted order.

All variables that have their highest absolute loadings on factor 1 are printed first, ordered from largest to smallest, then all variables that have their highest absolute loadings on factor 2 are printed, ordered from largest to smallest, and so on. This may make variables listed below any given factor appear unsorted. If K is specified then all loadings less than K are printed as 0.0.

SORT does not affect any storage commands.

BRIEF K

Controls the amount of printed output. BRIEF can be used as a main command as well as a subcommand. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output but performs all specified storage.
1	Minitab displays loadings (and sorted loadings if requested) for final solution.
2	In addition to the K = 1 output, Minitab displays factor score coefficients.
3	In addition to K = 2 output, Minitab displays information on the iterations when you use the ML subcommand.

CLUOBS: Session command for clustering observations

CLUOBS C...C

CLUOBS *M*

Performs agglomerative hierarchical clustering of observations when data are in columns (C...C) or on a distance matrix (M). The purpose of clustering observations is to help you classify observations into groups, when the groups are initially unknown.

Agglomerative clustering of observations begins with all observations separate, each forming its own cluster. In the first step, the two observations closest together are joined. In the next step, either a third observation joins the first two, or two other observations join together into a different cluster. Each step results in one less cluster than the step before until, at the end, all cases are combined in one cluster. Once two observations are combined

in a cluster, they may join with other observations, but they will always remain together. Minitab decides which clusters are closest together depending on how you define distance and which linking method you choose.

With CLUOBS you may enter either columns or a matrix as input arguments. Enter columns if you want to analyze raw data, the observations recorded for each variable. Enter a matrix if you have stored the distance matrix in M. The entry in row i and column j of distance matrix M is the distance between observations i and j.

Using subcommands, you can specify how to measure distance between observations, how to link observations together to form clusters, and what the final partition (final grouping of clusters) will be.

This command does not allow missing values.

Linkage method

These subcommands specify how Minitab determines which clusters are closest together, and thus which ones to combine. Use just one of them.

The linkage method you choose determines how the distance between two clusters is defined.

SINGLE
Specifies single linkage method.

AVERAGE
Specifies average linkage method.

COMPLETE
Specifies complete linkage method.

MCQUITTY
Specifies McQuitty linkage method.

CENTROID
Specifies centroid linkage method.

MEDIAN
Specifies median linkage method.

WARD
Specifies Ward linkage method.

Distance measure

If you do not supply a distance matrix, Minitab's first step is to calculate an N x N distance matrix, D, where N is the number of observations. Enter one of these subcommands to indicate which distance measure to use if you do not want to use the default, which is Euclidean. Specify just one. Do not use these subcommands if you list M as the argument with CLUOBS.

It is generally recommended that if you are using average, centroid, median, or Ward linkage, then you should use either squared Euclidean or squared Pearson distance. Euclidean is the standard mathematical measure of distance, Manhattan allows outliers to be down-weighted, and Pearson is a standardizing method.

EUCLIDEAN (default)
Specifies Euclidean distance measure.

SQEUCLIDEAN
Specifies squared Euclidean distance measure.

PEARSON
Specifies Pearson distance measure.

SQPEARSON
Specifies squared Pearson distance measure.

MANHATTAN
Specifies Manhattan distance measure.

Standardize

STANDARDIZE
Standardizes all variables (columns listed with CLUOBS) before the distance matrix is calculated. Specifically, it replaces a variable x with $(x - \bar{x})$ / standard deviation (x). Standardizing puts all variables on a common scale, so that one variable does not disproportionately influence the distance measure. If variables are in different units, standardizing them is a good practice.

Dendrogram

NUMBER K
Use NUMBER or CUT to specify the final partition. Use only one of these subcommands. If you do not use either of these subcommands, Minitab will continue clustering until all observations are combined into one cluster.

NUMBER says to cut the dendrogram so there are K clusters.

The output will contain information for the final partition. Distances printed here are always Euclidean distances. If STANDARDIZE is used, final partition characteristics are calculated using the standardized variables. Note: Final partition information cannot be calculated if the input is a distance matrix.

CUT K
Use NUMBER or CUT to specify the final partition. Use only one of these subcommands. If you do not use either of these subcommands, Minitab will continue clustering until all observations are combined into one cluster.

CUT says to cut the dendrogram at similarity level K.

The output will contain information for the final partition. Distances printed here are always Euclidean distances. If STANDARDIZE is used, final partition characteristics are calculated using the standardized variables. Note: Final partition information cannot be calculated if the input is a distance matrix.

DENDOGRAM
Produces a dendrogram or tree diagram, showing the amalgamation steps. For some data sets, average, centroid, median and Ward's methods do not produce a hierarchical dendrogram. That is, the amalgamation distances do not always increase with each step. In the dendrogram, such a step will produce a join that goes downward rather than upward.

DENDROGRAM has several subcommands of its own that allow you to customize the dendrogram.

TITLE *"text"*
Use TITLE to specify a title for the dendrogram.

CASE
Use CASE to specify case labels. The column of case labels must be the same length as the data column.

GSAVE *"file"*
GSAVE on page 39 allows you to save the dendrogram in a graph format. The default is a (PNG) file.

YSIMILARITY
Use YSIMILARITY or YDISTANCE to use either the similarity or distance on the y-axis.

YDISTANCE
Use YSIMILARITY or YDISTANCE to use either the similarity or distance on the y-axis.

PERPAGE K
Use PERPAGE to display K number of observations per graph. If you do not use PERPAGE, the dendrogram is displayed in a single graph.

Storage

DMATRIX *M*
Saves the N x N distance matrix, **D**, in M. M can be used in subsequent CLUOBS commands.

MEMBERSHIP C
Saves cluster membership for each observation in column C, provided either CUT or NUMBER was given. You can use C as a categorical variable in other Minitab commands, for example in DISCRIMINANT on page 922 or PLOT on page 1020.

DISTANCE C...C
Saves the distance between each observation and each cluster centroid, provided either CUT or NUMBER was given. The number of columns on DISTANCE must equal the number of clusters in the final partition. These distances are always Euclidean distances.

Results

BRIEF K
Controls the amount of output. K is an integer from 0 to 2. The larger the value of K, the more output. K = 2 is the default. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output but performs all specified storage.
1	Minitab displays the amalgamation steps and displays the dendrogram if you asked for it.
2 (default)	Minitab displays the K = 1 output, plus information on the final partition if either CUT or NUMBER is used.

CLUVARS: Session command for clustering variables

CLUVARS C...C

CLUVARS *M*

Performs agglomerative hierarchical clustering of variables. Input can be the data stored in columns C...C or a distance matrix stored in M.

The distance metric, the linkage type, and the final partition can be specified by subcommands.

Performs agglomerative hierarchical clustering of variables. Clustering of variables can help you classify variables into groups, when the groups are initially not known.

Agglomerative clustering of variables begins with all variables separate, each forming its own cluster. In the first step, the two variables closest together are joined. In the next step, either a third variable joins the first two, or two other variables join together into a different cluster. Each step results in one less cluster than the step before until, at the end, all variables are combined in one cluster. Once two variables are combined in a cluster, they may join with other variables, but they will always remain together. Minitab decides which clusters are closest together depending on how you define distance and which linking method you choose.

With CLUVARS you may enter either columns or a matrix as input arguments. Enter columns if you want to analyze raw data, the observations recorded for each variable. Enter a matrix if you have stored the distance matrix in M. The entry in row i and column j of M is the distance between variables i and j.

Using subcommands, you can specify how to measure distance between variables, how to link variables together to form clusters, and what the final partition (final grouping of variables) will be.

CLUVARS does not allow missing values.

Linkage method

These subcommands specify how Minitab determines which clusters are closest together, and thus which ones to combine. Use just one of them.

The linkage method you choose determines how the distance between two clusters is defined.

SINGLE

Specifies single linkage method.

AVERAGE

Specifies average linkage method.

COMPLETE

Specifies complete linkage method.

MCQUITTY

Specifies McQuitty linkage method.

CENTROID

Specifies centroid linkage method.

MEDIAN

Specifies median linkage method.

WARD

Specifies Ward linkage method.

Distance measure

These subcommands specify how to measure distance between clusters. Use one of these subcommands if the input to CLUVARS is columns (do not use if the input is a distance matrix). Specify only one. The default is CORRELATION.

If you did not supply a distance matrix, Minitab's calculates a P x P distance matrix using either the correlation or absolute correlation distance measure.

CORRELATION (default)
Specifies correlation to measure distance.

ABSCORRELATION
Specifies absolute correlation to measure distance.

Dendrogram

NUMBER K
Use NUMBER or CUT to specify the final partition. Use only one of these subcommands. If you do not use either of these subcommands, Minitab will continue clustering until all observations are combined into one cluster.

NUMBER says to cut the dendrogram so there are K clusters.

The output will contain information for the final partition. Distances printed here are always Euclidean distances. If STANDARDIZE is used, final partition characteristics are calculated using the standardized variables. Note: Final partition information cannot be calculated if the input is a distance matrix.

CUT K
Use NUMBER or CUT to specify the final partition. Use only one of these subcommands. If you do not use either of these subcommands, Minitab will continue clustering until all observations are combined into one cluster.

CUT says to cut the dendrogram at similarity level K.

The output will contain information for the final partition. Distances printed here are always Euclidean distances. If STANDARDIZE is used, final partition characteristics are calculated using the standardized variables. Note: Final partition information cannot be calculated if the input is a distance matrix.

DENDOGRAM
Produces a dendrogram or tree diagram, showing the amalgamation steps. For some data sets, average, centroid, median and Ward's methods do not produce a hierarchical dendrogram. That is, the amalgamation distances do not always increase with each step. In the dendrogram, such a step will produce a join that goes downward rather than upward.

DENDROGRAM has several subcommands of its own that allow you to customize the dendrogram.

TITLE *"text"*
Use TITLE to specify a title for the dendrogram.

GSAVE *"file"*
GSAVE on page 39 allows you to save the dendrogram in a graph format. The default is a (PNG) file.

YSIMILARITY
Use YSIMILARITY or YDISTANCE to use either the similarity or distance on the y-axis.

YDISTANCE
Use YSIMILARITY or YDISTANCE to use either the similarity or distance on the y-axis.

PERPAGE K

Use PERPAGE to display K number of observations per graph. If you do not use PERPAGE, the dendrogram is displayed in a single graph.

Storage

DMATRIX *M*

Saves the N x N distance matrix, **D**, in M. M can be used in subsequent CLUOBS commands.

Results

BRIEF K

Controls the amount of output. K is an integer from 0 to 2. The larger the value of K, the more output. K = 2 is the default. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output but performs all specified storage.
1	Minitab displays the amalgamation steps and displays the dendrogram if you asked for it.
2 (default)	Minitab displays the K = 1 output, plus information on the final partition if either CUT or NUMBER is used.

KMEANS: Session command for non-hierarchical clustering of observations

KMEANS C...C

Performs K-means non-hierarchical clustering of observations according to MacQueen's algorithm.

Like CLUOBS on page 915, the purpose of K-means clustering is to classify observations into groups when the groups are initially unknown. KMEANS begins with a grouping of observations into a pre-defined number of clusters. Minitab evaluates each observation, moving it into the cluster whose centroid it is closest to, using Euclidean distance. When a cluster changes, by losing or gaining an observation, Minitab recalculates the cluster centroid. This process is repeated until no more observations can be moved into a different cluster, that is, until all observations are in the cluster whose centroid they are closest to.

In CLUOBS, which is a hierarchical procedure, once two observations are joined in a cluster, they can never be separated. In K-means clustering, which is not hierarchical, there is no such constraint. KMEANS uses Euclidean distances to measure distance between observations and clusters.

With KMEANS, list the columns containing the data. You must use either the NUMBER or INITIAL subcommand.

KMEANS does not allow missing values.

STANDARDIZE

This subcommand standardizes all variables (columns listed with KMEANS) before any calculations are done. Standardizing puts all variables on a common scale, so that one variable does not disproportionately influence the distance measure. If variables are in different units, it is a good idea to standardize them.

If STANDARDIZE is used, the final cluster centroids and cluster characteristics are printed in the standardized variable space.

MEMBERSHIP C

Stores final cluster membership for each observation in C. You can use C as a categorical variable in other Minitab commands, for example in DISCRIMINANT on page 922 or PLOT on page 1020.

DISTANCE C...C

Saves the distance between each observation and each cluster centroid. DISTANCE must list the same number of columns as the number of clusters requested with NUMBER or INITIAL. These are Euclidean distances.

BRIEF K

Controls the amount of output. BRIEF 0 suppresses all output, but does all requested storage. BRIEF 2, the default, displays all output. BRIEF can be used as a main command as well as a subcommand.

Initial partition for the KMEANS algorithm

You must use one of these subcommands to specify the initial partition for the KMEANS algorithm.

The final partition will depend to some extent on the initial partition that Minitab uses. You can use INITIAL to try different initial partitions.

NUMBER K

NUMBER specifies how many clusters to form. The first K observations are taken as initial cluster centroids. Each observation is assigned to the cluster whose centroid it is closest to. Minitab recalculates the cluster centroids each time a cluster gains or loses an observation.

For example, NUMBER 3 says to use the first three observations as the initial cluster centroids, and to go through the K-means process resulting in a final partition of three clusters.

INITIAL C

INITIAL specifies the initial partition. C must be the same length as the input columns and contain positive, consecutive integers or zeros (it should not contain all zeros). Initially, each observation is assigned to the cluster identified by the corresponding value in C. Observations corresponding to a value of zero in C are initially unassigned. The number of distinct positive integers in C equals the number of clusters in the final partition.

DISCRIMINANT: Session command for performing discriminant analysis

DISCRIMINANT C, C...C

Performs linear and quadratic discriminant analysis for classifying observations into two or more groups based on the specified predictors, with groups in C and predictors in C...C.

In linear discriminant analysis, all groups are assumed to have the same covariance matrix. Quadratic discrimination does not make this assumption. Its properties, however, are not as well understood.

The column of group codes may be numeric or text, and may contain any values. The codes do not need to be in any special order. When the group column is text, the first entry (row 1) is level 1, the next entry that is different is level 2, and so on.

Output includes the classification matrix (sometimes called the confusion matrix), the squared distance between group centers, the linear discriminant function (unless you requested quadratic discrimination), means, standard deviations, covariance matrices, and a summary of how each observation was classified. You can use cross-validation to examine the goodness of classification.

QUADRATIC

Requests quadratic discrimination. The default is linear. QUADRATIC cannot be used with LDF.

PRIORS K...K

Specifies prior probabilities. There should be one value for each group. The first probability will be assigned to the group with the smallest code, the second to the group with the second smallest code, etc. If the probabilities do not sum to one, Minitab normalizes them.

XVAL

Performs discrimination using cross-validation. This procedure omits the first observation from the data set, develops a classification function using the remaining observations, then classifies the omitted observation. Next, it returns the first observation to the data set, omits the second observation, and repeats the same process. XVAL continues in this manner with all observations in the data set.

XVAL prints the classification matrix using cross-validation, and a summary of how each observation was classified. Using XVAL increases computation time by approximately a factor of four.

XVAL is one technique used to compensate for an optimistic apparent error rate (AER). AER is (number of misclassified observations in the data set)/(total number of observations in the data set). AER tends to be optimistic because the data being classified are the same data used to build the classification function.

Another technique to calculate a more realistic error rate is to split your data set into two parts. Use one part as a training set on which to build the discriminant function, and the other part as a validation set on which to evaluate the function (use the PREDICT subcommand).

LDF C...C

Stores the coefficients from the linear discriminant function, using one column for each group. The constant is stored at the top of the column. LDF cannot be used with QUADRATIC.

FITS C [C]

Stores the fitted values. The fitted value for an observation is the group into which it is classified. The optional second column can be used to store the fitted values from cross-validation (provided you have used the subcommand XVAL).

PREDICT *E...E*

Specifies the predicted group membership for new observations in E...E. A summary table is displayed. The number of columns or constants must be equal to the number of predictors.

If you execute DISCRIMINANT from the menu, only one PREDICT subcommand is allowed.

BRIEF K

Controls the amount of printed output. BRIEF can be used as a main command as well as a subcommand, but is not available if you execute DISCRIMINANT from the menu.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the classification matrix.
2 (default)	In addition to the K = 1 output, Minitab displays the distance between all pairs of group centers (i.e., group means), the linear discriminant function, and a summary of misclassified observations.
3	In addition to K = 2 output, Minitab displays the means, standard deviations, and covariance matrices (for each group and pooled).
4	In addition to K = 3 output, Minitab displays a summary of how all observations were classified. Minitab marks misclassified observations with two asterisks.

ITEMANALYSIS: Session command for performing item analysis

ITEMANALYSIS C...C

Calculates Cronbach's alpha and basic statistics for items or variables in C...C.

Use item analysis to assess whether the items in a survey or test are measuring the same construct. You must have two or more numeric columns. Typically, the data will be a score of some type, such as test scores or Likert-scale responses. Each column must be of equal length. Any row with a missing value is not included when calculating the output.

Options

STANDARDIZE

Standardizes the data before calculating the Cronbach's alpha and the omitted item statistics. Standardizing puts all variables on a common scale, so that one or a few variables do not disproportionately influence the results.

Graphs

GMATRIX

Plots a matrix plot that contains a scatterplot of every unique combination of items in the analysis.

Results

RCORRELATION

Displays the correlation matrix.

RCOVARIANCE

Displays the covariance matrix.

RSUMMARY

Displays the summary statistics of item and the total.

RALPHA
Displays Cronbach's alpha.

ROMIT
Displays adjusted statistics by omitting items one at a time. Displayed statistics includes, mean, standard deviation, correlation between the omitted item and the total of remaining items, multiple correlation when the omitted item is regressed on the remaining items, and Cronbach's alpha.

CA: Session command for performing simple correspondence analysis

CA C...C
Simple correspondence analysis helps you explore relationships in a two-way contingency table, that is, a table of frequencies based on two categorical variables. You can also use simple correspondence analysis to analyze three-way and four-way contingency tables. See the RCROSS and CCROSS subcommands for details.

Input data

CONTINGENCY
Specifies the arguments listed on the main command line as columns of a contingency table. When you do not use CONTINGENCY, Minitab assumes the input columns are categorical variables. All values in the contingency columns must be positive integers or zero.

RNAMES C
Specifies row names for the contingency table.

The name column must be a text column whose length matches the number of rows or columns in the contingency table. CA displays the first 8 characters of the names in tables, but displays the full name on graphs.

CNAMES C
Specifies column names for the contingency table.

The name column must be a text column whose length matches the number of rows or columns in the contingency table. CA prints the first 8 characters of the names in tables, but prints the full name on graphs.

NCOMPONENTS K
Specifies the number of components to calculate. The minimum number of components is one. The maximum number of components for a contingency table with r rows and c columns is the smaller of (r-1) or (c-1). This is equivalent to the dimension of the subspace onto which you project the profiles. The default number of components is 2.

Combine

Cross two variable columns to form the rows and/or columns of the contingency table allowing you to use simple correspondence analysis to analyze three and four-way contingency tables.

Both RCROSS and CCROSS require that you enter your data as categorical variables rather than a contingency table. If you use just RCROSS or just CCROSS, you must specify 3 columns on the main command line. If you use both RCROSS and CCROSS, you must specify 4 columns on the main command line.

RCROSS crosses the categories in the first two input columns to form the rows of the contingency table. For example, if the first variable is Sex (with 2 levels, male and female) and the second variable is Age (with 3 levels, young, middle aged, old), then there will be 2 x 3 = 6 rows, ordered as follows:

males / young

males / middle aged

males / old

females / young

females / middle aged

females / old

CCROSS crosses the categories in the last two input columns to form the columns of the contingency table.

RCROSS
Crosses the first two variables to form the rows.

CCROSS
Crosses the last two variables to form the columns.

Supplementary data

RSUPPLEMENTARY AND CSUPPLEMENTARY specify supplementary data for the analysis. CA does not use the supplementary points in calculating the components (i.e., to determine the reduced dimensions), but displays table information on supplementary points. To include supplementary points in the plots, use the SUPPLEMENTARY subcommand. Use RSUPPLEMENTARY to specify supplementary rows. List one column for each supplementary row. There is just one form for the supplementary data—each worksheet column is one row of the contingency table. If there are c columns in the contingency table, there will be c entries in each worksheet column. Use CSUPPLEMENTARY to specify supplementary columns. List one worksheet column for each supplementary column. There is just one form for the supplementary data—each worksheet column is one column of the contingency table. If there are r rows in the contingency table, there will be r entries in each worksheet column.

SRNAMES and SCNAMES give names for the supplementary rows and supplementary columns of the contingency table, respectively. The specified name column must be a text column whose length matches the number of supplementary rows or supplementary columns in the contingency table. CA prints the first eight characters of the names in tables, but prints the full name on graphs.

RSUPPLEMENTARY C...C
Specifies row supplementary data.

CSUPPLEMENTARY C...C
Specifies column supplementary data.

SRNAMES C
Specifies names for the supplementary rows.

SCNAMES C
Specifies names for the supplementary columns.

Results

CTABLE

CTABLE displays the contingency table. The entry in row i and column j of the contingency table is the frequency for that cell. The total in row i is the sum of the frequencies in row i. The total in column j is the sum of the frequencies in column j.

RPROFILES

RPROFILES displays a table of row profiles and row masses.

CPROFILES

CPROFILES displays a table of column profiles and column masses.

EXPECTEDS

EXPECTEDS displays a table of the expected frequency in each cell of the contingency table.

DIFFERENCES

DIFFERENCES displays a table of the observed minus the expected frequency in each cell of the contingency table.

CHISQUARE

CHISQUARE displays a table of the χ^2 value in each cell of the contingency table.

INERTIAS

INERTIAS displays a table of the relative inertia in each cell of the contingency table.

Graphs

RPLOT

RPLOT displays a plot that shows the row principal coordinates. Row points are plotted with red circles: solid red circles for regular points, and open red circles for supplementary points.

CPLOT

CPLOT displays a plot that shows the column principal coordinates. Column points are plotted with blue squares: solid blue squares for regular points, and open blue squares for supplementary points.

SPLOT

Displays a symmetric plot that shows both row and column coordinates overlaid in a joint display. CA plots points that represent rows with red circles: solid red circles for regular points, and open red circles for supplementary points. CA plots points that represent columns with blue squares: solid blue squares for regular points, and open blue squares for supplementary points.

ARPLOT

ARPLOT displays an asymmetric row plot.

ACPLOT

ACPLOT displays an asymmetric column plot. Points that represent rows are plotted with red circles.

AXES K K, ..., K K

Specifies the axis pairs for each requested plot. You may list between 1 and 15 pairs. The axes you list must be axes in the subspace you defined on NCOMPONENTS. For example, if you specified `NCOMPONENTS 4` then you can only use axes 1, 2, 3, and 4.

The first axis in a pair will be the Y or vertical axis of the plot; the second axis will be the X or horizontal axis of the plot. For example, AXES 2 1 3 1 plots component 2 versus component 1, and component 3 versus component 1.

SUPPLEMENTARY

Includes supplementary points on all plots. CA plots points that represent rows with red circles: solid red circles for regular points, and open red circles for supplementary points. CA plots points that represent columns with blue squares: solid blue squares for regular points, and open blue squares for supplementary points.

Asymmetric row plot and asymmetric column plot

ARPLOT displays an asymmetric row plot. ACPLOT displays an asymmetric column plot. Points that represent rows are plotted with red circles. CA plots points that represent rows with red circles: solid red circles for regular points, and open red circles for supplementary points. CA plots points that represent columns with blue squares: solid blue squares for regular points, and open blue squares for supplementary points.

A disadvantage of asymmetric plots is that the profiles of interest are often bunched in the middle of the graph. A symmetric plot of both rows and columns spreads out both profiles. However, you must be careful interpreting the symmetric plot because this plot overlays two different mappings.

ARPLOT

ARPLOT displays an asymmetric row plot.

An asymmetric row plot shows both the coordinates of the row profiles projected onto the low dimensional subspace (these are the coordinates that the subcommand RPCOORDINATES stores), and the coordinates of the column vertices projected onto the low dimensional subspace (these are the coordinates that the subcommand CSCOORDINATES stores). The vertex corresponding to column j is the point in row space with a 1 in position j and 0's elsewhere, i.e., it is the row profile which is totally concentrated in one column group.

ACPLOT

ACPLOT displays an asymmetric column plot. Points that represent rows are plotted with red circles.

An asymmetric column plot shows both the coordinates of the column profiles projected onto the low dimensional subspace (these are the coordinates that the CPCOORDINATES subcommand stores), and the coordinates of the row vertices projected onto the low dimensional subspace (these are the coordinates that the RSCOORDINATES subcommand stores). The vertex corresponding to row i is the point in column space with a 1 in position i and 0's elsewhere, i.e., it is the column profile which is totally concentrated in one row group.

Storage

RPCOORDINATES C...C

RPCOORDINATES stores the row principal coordinates. CA stores the coordinate for the first component in the first column, the coordinate for the second component in the second column, etc. If there are supplementary points, their coordinates are stored at the ends of the columns.

The principal coordinate for row profile i and component (axis) k is the coordinate of the projection of row profile i onto axis k. These are the points that are plotted when you ask for a row plot (RPLOT).

The row standardized coordinates for component k are the principal coordinates for component k divided by the square root of the k^{th} inertia.

RSCOORDINATES C...C

RSCOORDINATES stores the row standardized coordinates. CA stores the coordinate for the first component in the first column, the coordinate for the second component in the second column, etc. If there are supplementary points, their coordinates are stored at the ends of the columns.

The principal coordinate for row profile i and component (axis) k is the coordinate of the projection of row profile i onto axis k. These are the points that are plotted when you ask for a row plot (RPLOT).

The row standardized coordinates for component k are the principal coordinates for component k divided by the square root of the k^{th} inertia.

CPCOORDINATES C...C

CPCOORDINATES stores the column principal coordinates. The coordinate for the first component is stored in the first column, the coordinate for the second component in the second column, etc. If there are supplementary points, their coordinates are stored at the ends of the columns.

The principal coordinate for column profile j and component (axis) k is the coordinate of the projection of column profile j onto axis k. These are the points that are plotted when you ask for a column plot (CPLOT).

The column standardized coordinates for component k are the column principal coordinates for component k divided by the square root of the k^{th} inertia.

CSCOORDINATES C...C

CSCOORDINATES stores the column standardized coordinates. The coordinate for the first component is stored in the first column, the coordinate for the second component in the second column, etc. If there are supplementary points, their coordinates are stored at the ends of the columns.

The principal coordinate for column profile j and component (axis) k is the coordinate of the projection of column profile j onto axis k. These are the points that are plotted when you ask for a column plot (CPLOT).

The column standardized coordinates for component k are the column principal coordinates for component k divided by the square root of the k^{th} inertia.

SVTB C...C

Stores the contingency table. List one worksheet column for each column of the contingency table. SVTB does not save supplementary rows and columns.

MCA: Session command for performing a multiple correspondence analysis

MCA C...C

Multiple correspondence analysis extends simple correspondence analysis to the case of three or more categorical variables.

There are two forms for the input columns listed on the main command line: categorical or indicator variables. If you input the data as indicator variables, you do need to use the INDICATOR subcommand. This command does not accept data with missing values. Remove any cases with missing data before you use this command.

Input data

INDICATORS

Specifies the input data as indicator variables (also called a disjunctive matrix). List one worksheet column for each level of each categorical variable. The entries in all columns must be the integers 0 and 1. All columns must be the same length and may not contain missing values.

CNAMES C

Gives the category names. The name column must be a text column whose length matches the number of categories on all categorical variables.

For example, suppose there are 3 categorical variables: Sex (male, female), Hair color (blond, brown, black), and Age (under 20, from 20 to 50, over 50), and no supplementary variables. You would assign 2 + 3 + 3 = 8 category names, and the CNAMES column would contain 8 rows.

MCA only uses the first 8 characters of the names in displayed tables, but uses all characters on graphs.

NCOMPONENTS K

Specifies the number of components to calculate. The default number of components is 2.

Supplementary data

CSUPPLEMENTARY C...C

Specifies supplementary data for the analysis. MCA does not use the supplementary points in calculating the components (i.e., to determine the reduced dimensions), but prints table information on supplementary points. To include supplementary points in the plots, use the SUPPLEMENTARY subcommand.

Specify supplementary data using the same form, either categorical variables or indicator variables, as you did for the data specified on the main command line.

SCNAMES C

Specifies names for the supplementary data. The name column must be a text column whose length matches the number of categories for all supplementary variables. Only the first 8 characters of the names will be used in printed tables; all characters will be used on graphs. If you do not enter names here, the rows will be named Supcol1, Supcol2, and so on.

Results

ITABLE

Displays the indicator table. There is one indicator variable (column of the worksheet) for each level (category) of each categorical variable. An observation (row) has a 1 in columns that correspond to levels that are true for this observation and a 0 otherwise. An indicator table (also called a disjunctive matrix) has one column for each level (category) of each categorical variable.

BURT

Displays the Burt table. A Burt matrix is a symmetric matrix which has one column and one row for each level (category) of each categorical variable. The entry in row i and column j is the number of observations that have the level corresponding to row i and also the level corresponding to column j.

Graphs

AXES K K, ..., K K

Specifies the axis pairs for the column plot. You may list between 1 and 15 pairs. The axes you list must be axes in the subspace you defined on NCOMPONENTS. For example, if you specified `NCOMPONENTS 4`, then you can use only axes 1, 2, 3, and 4.

The first axis in a pair will be the Y or vertical axis of the plot; the second axis will be the X or horizontal axis of the plot. For example, `AXES 2 1 3 1` plots component 2 versus component 1, and component 3 versus component 1.

SUPPLEMENTARY

Includes supplementary points in the column plot. MCA plots regular points as solid blue squares and supplementary points as open blue squares.

CPLOT

Displays a plot that shows the column coordinates. Column points are plotted with blue squares - solid squares for regular points, and open squares for supplementary points. To include supplementary points in the column plot, use the SUPPLEMENTARY subcommand.

Storage

COORDINATES C...C

Stores the column coordinates. MCA stores the coordinate for the first component in the first listed column, the coordinate for the second component in the second listed column, and so on. If there are supplementary points, their coordinates are stored at the ends of the columns.

Time Series Analysis

AAMODEL: Session command for the selection of an alternative model from **Forecast with Best ARIMA Model**

AAMODEL K K

After the analysis identifies the best model with ATARIMA, AAMODEL produces results for alternative models. You can use AAMODEL only when the ATARIMA command used the SMODEL subcommand to create the object ID and model ID.

The first K is a positive integer that represents a **Forecast with Best ARIMA Model** object ID. The first **Forecast with Best ARIMA Model** that you create in a project has the object ID 1. The second **Forecast with Best ARIMA Model** that you create with the command has the object ID 2, and so on.

The second K represents the model ID. The value of the second K is an integer from 2 to 30. The model ID refers to the model with that row number in the Model Selection table. The Model Selection table is for the object ID that you specify with the first K.

The analysis displays the tables and graphs in a new output tab. The same columns that the original analysis stored are also stored for the new model. The storage for the new model is in a new column. For example, if the original analysis stored the fitted values in a column titled "Fit," then the new analysis names an empty column "Fit_1" and stores the fitted values in that column.

ACF: Session command for calculating autocorrelation

ACF C

Calculates and plots the autocorrelations of a time series.

ACF generates a high-resolution autocorrelation function, including confidence limits for the correlations, in a graph. You can store the autocorrelations, associated Ljung-Box Q statistics, and t-statistics for up to the first 75 lags.

If you do not specify the number of lags with LAGS, then Minitab uses a default number of lags:

- n / 4 for a series with less than or equal to 240 observations
- $\sqrt{n}$ + 45 for a series with more than 240 observations, where n is the number of observations in the series

LAGS K

Specifies the number of lags to display in the ACF plot. If you omit LAGS, then Minitab uses a default number of lags. This is n / 4 for a series that has less than or equal to 240 observations, or $\sqrt{n}$ + 45 for a series that has more than 240 observations, where n is the number of observations in the series.

SACF C
Stores the autocorrelations.

TSTATS C
Stores the t-statistics.

LBQSTATS C
Stores the Ljung-Box Q statistics.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

ADF: Session command for conducting an augmented Dickey-Fuller test

ADF C
ADF performs an augmented Dickey-Fuller test on time series data to determine whether differencing a series makes the mean of the data stationary. Usually, you use the augmented Dickey-Fuller test to determine the nonseasonal differencing order when you analyze your time series data with an ARIMA model.

MAXLAG K
Specifies the maximum lag order to evaluate when fitting the regression model to calculate the test statistic. K must be an integer in the interval [0, floor(n/2)-3] where n is the row count of the column for ADF, omitting any missing values from the beginning or end of the data. The default value if you do not specify this subcommand is min{floor(n/2 - 3), floor[12 * (n/100)^(1/4)]}.

AUTOLAG K K
Specifies the criterion to use to select the lag order in the regression model. The default value if you do not specify this command is 1. The analysis evaluates the regression model with the maximum lag order first, then removes one lag at a time, beginning with the highest lag order.

K	Criterion
0	The analysis uses all lag orders less than or equal to the maximum lag order to fit the regression model.
1	The selected lag order in the regression model minimizes the Akaike Information Criterion (AIC).
2	The selected lag order in the regression model minimizes the Bayesian Information Criterion (BIC).
3	When the t-statistic for the highest lag order is significant, the analysis selects that order for the regression model.

ATERMS K
Specifies additional terms other than the lag terms to include in the regression model. The default value if you do not specify this command is 1 so that the mean of the series is non-zero when the series is stationary.

K	Additional terms
0	Include no additional terms in the model so that the mean of the series is 0 if the series is stationary.
1	Include a constant term in the model.
2	Include a constant term and a linear term in the model.
3	Include a constant term, a linear term, and a quadratic term in the model.

SLEVEL K

Specifies the significance level for the test. K must be in the interval (0.000001, 0.999999). The default value if you do not specify this subcommand is 0.05.

If the significance level is 0.01, 0.05, or 0.10 then the recommendation of the analysis depends on a comparison of the test statistic and the critical value. If the test statistic is less than or equal to the critical value, reject the null hypothesis that the data are non-stationary. Rejection of the null hypothesis leads to the recommendation that differencing is unnecessary.

If the significance level is any other value, then the recommendation of the analysis depends on a comparison of the p-value to the significance level. If the p-value is less than or equal to the significance level, reject the null hypothesis that the data are non-stationary. Rejection of the null hypothesis leads to the recommendation that differencing is unnecessary.

TMETHOD

Displays the method table.

TTEST

Displays table of the results of the augmented Dickey-Fuller test.

GPLOT

Displays a time series plot, Autocorrelation Function (ACF) plot, and Partial Autocorrelation Function (PACF) plot of the original time series data.

GDPLOT

Displays a time series plot, Autocorrelation Function (ACF) plot, and Partial Autocorrelation Function (PACF) plot of the differenced time series data.

ARIMA: Session command for modeling time series behavior and generating forecasts

ARIMA K K K K K K K C [C [C [C]]]

The ARIMA command fits nonseasonal and seasonal models to a time series.

The seven Ks specify the model, using the usual notation: (pdq) x (PDQ) S

Thus, the first K = p, the second K = d, the third K = q. If you have a nonseasonal model, that's all you need. If you have a seasonal model, then use the remaining four Ks to specify P, D, Q, and S.

The first column is for the data. The three optional columns are for storage. If you specify one column, then Minitab stores the residuals. If you specify two columns, then Minitab also stores the predicted values. If you specify a third column, then Minitab also stores the estimated parameters, in the same order they are displayed on the output.

The maximum order of each term, p, q, P, Q, is 5. The maximum number of parameters is 10. The maximum back order is 100. That is, if the model is expanded out, then the highest exponent for the back-shift operator must be < 100. At least one of the parameters p, q, P, Q must be non-zero. There must be at least 3 (nonmissing) values left after differencing.

The BRIEF subcommand can be used to control the amount of output that is produced.

CONSTANT

CONSTANT specifies to include a constant term in the model. If d = 0 and D = 0 (or is not used), this is the default.

NOCONSTANT

NOCONSTANT specifies to not include the constant term in the model. If d > 0 or D > 0, this is the default.

STARTING C

Specifies starting values (initial values) for the parameters. The values are entered in the order that the parameters appear on the output: ARs (autoregressive parameters), seasonal ARs, MAs (moving average parameters), seasonal MAs, and constant. The starting value for the constant is optional.

If you do not use START, Minitab uses 0.1 for all the parameters except the constant. This usually works well.

ARIMA performs up to 25 iterations to fit a given model. If you require more than 25 iterations to get a good fit, store the estimated parameters. Then use them as STARTING values for a second fit. This can be repeated as often as desired.

FORECAST [K] K [C [C C]]

Allows you to forecast observations starting at the specified origin and going up to K leads ahead. You can specify up to 150 leads. If you do not specify the origin, then the origin is the end of the series and the forecasts are for the future. Store the forecasts in C and specify the storage columns C C for the lower and upper confidence limits.

Suppose that you have yearly data from 1980 to 2018. The subcommand `FORECAST 5` forecasts the next 5 years, 2019 to 2023. The subcommand `FORECAST 31 5` forecasts from the 31st observation, 2010, for the next 5 years, 2011 to 2015.

Storing forecasts and confidence limits allows a variety of plots.

GSERIES

Displays a time series plot of the series. When you use the FORECASTS subcommand, Minitab displays the forecasts and their 95% confidence limits on the plot.

GACF

GACF displays an autocorrelation function for the standard or raw residuals.

GACF and GPACF include 95% confidence limits for the standard or raw residuals.

GPACF

GPACF displays a partial autocorrelation function for the residuals.

GACF and GPACF include 95% confidence limits for the standard or raw residuals.

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE
Displays a separate graph for the residuals versus each specified column.

BRIEF K
Controls the amount of output. You can also use BRIEF as a main command. When you use BRIEF as a main command, it affects all other commands that use BRIEF to control the amount of output.

Value of K	Output that is displayed
0	Minitab displays no output, and does not store any statistics
1	Minitab displays tables of final estimates, difference information, residual sum of squares, and number of observations
2 (default)	Same as K = 1 output, and table of estimates for each iteration
3	Same as K = 2 output, and correlation matrix
4	Same as K = 3 output, and back forecasts

ATARIMA: Session command for selecting parameters for an ARIMA model and generating forecasts

ATARIMA C
Evaluates specified combinations of terms in ARIMA models to identify the best model for the time series data in column C. Candidate ARIMA models with a specified order of differencing (d) can include different orders of autoregressive (p), and moving average terms (q). Candidate ARIMA models can also include a specified order of seasonal differencing (D) and different orders of seasonal terms (P, Q). To evaluate non-seasonal models, use the `BMODEL` subcommand. To evaluate seasonal models, use the `BMODEL` subcommand and the `SEASON` subcommand.

The analysis allows missing values at the beginning or end of the column. The analysis does not allow missing values to separate numbers in the column.

The analysis uses a specified criterion to determine the best model for the combination of the orders of the autoregressive and moving average terms in the search. The specifications for the analysis provide the orders of non-seasonal and seasonal differencing. The analysis provides results for the best ARIMA model.

BMODEL K K K K K

Specify the order of the nonseasonal terms to evaluate. The first K specifies the nonseasonal order of differencing, d. The last 4 Ks specify the range of orders to evaluate for the autoregressive terms, p, and the moving average terms, q.

Constant	Order
First K	Nonseasonal differencing order
Second K	Minimum autoregressive order
Third K	Maximum autoregressive order
Fourth K	Minimum moving average order
Fifth K	Maximum moving average order

The order of differencing is 0, 1, or 2. For autoregressive and moving average orders, enter values in the interval [0, 5]. For the pairs of constants that define the range for the autoregressive and the moving average orders, enter a minimum order ≤ the maximum order. If the two values are equal, then all of the candidate models use that order.

SEASON K K K K K K K

Specify the seasonal period and the order of the seasonal terms to evaluate. The first K specifies the seasonal period. The second K specifies the seasonal differencing order. The last 4 Ks specify the range of seasonal orders to evaluate for the autoregressive terms, P, and the moving average terms, Q.

Constant	Order
First K	Seasonal period
Second K	Seasonal differencing order
Third K	Minimum seasonal autoregressive order
Fourth K	Maximum seasonal autoregressive order
Fifth K	Minimum seasonal moving average order
Sixth K	Maximum seasonal moving average order

For the seasonal period, enter a value in the interval [2, FLOOR(N/2)]. In other words, seasonal periods are an integer greater than 1 and less than half of the length of the series.

The order of differencing is 0, 1, or 2. For seasonal autoregressive and seasonal moving average orders, enter values in the interval [0, 5]. For the pairs of constants that define the range for the seasonal autoregressive and the seasonal moving average orders, enter a minimum order ≤ the maximum order. If the two values are equal, then all of the candidate models use that order.

CONSTANT

Specifies to include a constant term in candidate models that have d + D ≤ 1 where d is the non-seasonal order of differencing and D is the seasonal order of differencing. If the parameter estimates for a candidate model with a constant do not converge, the analysis evaluates the model without the constant term. If you do not use this subcommand, then the analysis fits all the candidate models without a constant term.

Requirements for model orders

For a candidate model with a constant, the analysis requires that p + q + P + Q ≤ 9. For a model without a constant, the analysis requires that p + q + P + Q ≤ 10. These requirements mean that, for example, the ARIMA model (5, 0, 5) with a constant is not a candidate model even if the analysis includes BMODEL 0 0 5 0 5.

For a candidate model, the analysis requires that at least 1 of p, q, P, or Q > 0. The only exception is the ARIMA model (0, 1, 0). You can fit the ARIMA model (0, 1, 0) with or without a constant. The ARIMA model (0, 1, 0) is also known as the random walk model.

Criteria

The analysis uses Akaike's Corrected Information Criterion (AICc), Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC) to select the best model among models with the same order of differencing. If you do not specify which criterion to use, the analysis uses AICc.

AICC
Specifies the use of the corrected Akaike's Information Criterion (AICc) to select the best model.

AIC
Specifies the use of the Akaike's Information Criterion (AIC) to select the best model.

BIC
Specifies the use of the Bayesian Information Criterion (BIC) to select the best model.

Forecasts

FORECAST [K] K
Allows you to forecast observations starting at the specified origin and going up to K leads ahead.

The optional first K specifies the origin. If you do not specify the origin, then the origin is the end of the series and the forecasts are for the future. If you specify the origin, the index for K begins at the first non-missing value in the series. The second K specifies the number of forecasts. You can specify up to 150 forecasts.

Suppose that you have yearly data from 1980 to 2018. The subcommand `FORECAST 5` forecasts the next 5 years, 2019 to 2023. The subcommand `FORECAST 31 5` forecasts from the 31st observation, 2010, for the next 5 years, 2011 to 2015.

CONFIDENCE K
Specifies a confidence level for the lower and upper probability limits of the forecasts. K is in the interval (50, 100). For example, for a 90% confidence level, enter `CONFIDENCE 90`. If you do not use this subsubcommand, then the confidence level is 95.

Box-Cox

BOXCOX [K]
Performs a Box-Cox transformation with a specified lambda. K is the value of lambda and must be between −5 and +5. If you do not specify K, then Minitab finds the optimal lambda. By default, Minitab rounds the optimal value.

LULIMITS K K
Specifies the lower and upper limits for the search for the optimal lambda. The first K is the lower limit and must be less than the second K. Both values of K must be within [-5, 5]. The default values if you do not specify `LULIMITS` are -1, 2.

Graphs

GTSERIES [C]

Displays a time series plot of the series. When you use the FORECAST subcommand, Minitab displays the forecasts and their 95% confidence limits on the plot. C is an optional column that gives tick labels for the X-axis. The column C must have the same length as the column for the main command. The column with the tick labels can have any data type. If you use the BOXCOX command, then the GTSERIES command produces plots for both the original data and the transformed data.

GACF

Displays an autocorrelation function for the regular residuals. The plot includes confidence limits.

GPACF

Displays a partial autocorrelation function for the regular residuals. The plot includes confidence limits.

RTYPE K

Specifies the type of residual to plot with following subcommands: GHISTOGRAM, GNORMAL, GFITS, GORDER, GFOURPACK, GVARIABLE.

Value of K	Type of residuals
1	Regular residuals
2 (default)	Standardized residuals

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C

Displays a separate graph for the residuals versus each specified column.

Results

NODEFAULT

Specifies that no default tables or graphs will be displayed

TMETHOD

Displays the method table.

TSELECTION
Displays the model selection table.

TITERATION
Displays the table that shows the parameter estimates for each iteration that leads to the best model.

TBFORECAST
Displays the table of back forecasts for the best model.

TBRESIDUAL
Displays the table of back forecast residuals for the best model.

TPARAMETER
Displays the final estimates of the parameters for the best model.

TCORRELATION
Displays the correlation matrix for the estimated parameters for the best model.

TSUMMARY
Displays the model summary table for the best model.

TCHISQUARE
Displays the modified Box-Pierce (Ljung-Box) chi-square statistic for the best model.

TFORECAST
Displays the forecast table for the best model.

Storage

SMODEL
Stores IDs for the alternative models so that you can use the `AAMODEL` command to produce ARIMA results for the models in rows 2 to 30 of the Model Selection table.

SFITS C
Stores the fitted values. When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SRESIDUALS C
Stores the residuals (fitted values – observed values). When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SSRESIDUALS C
Stores the standardized residuals. When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SCOEF C
Stores the coefficients from the best model.

SFORECAST C
Stores the forecast values from the best model. When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SSEFORECAST C
Stores the standard errors of the forecast values. When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SLULLIMITS C C
Stores the lower and upper probability limits for the forecast values. When the analysis uses a Box-Cox transformation, these values are in the transformed units.

SBCSERIES C
Stores the Box-Cox transformed data when the analysis uses a Box-Cox transformation.

SOFIT C
Stores the fitted values from the best model for the original series when the analysis uses a Box-Cox transformation.

SOFORECAST C
Stores the forecast values from the best model for the original series when the analysis uses a Box-Cox transformation.

SOLULIMITS C
Stores the lower and upper probability limits for the forecast values for the original series when the analysis uses a Box-Cox transformation.

CCF: Session command for calculating cross correlation between two time series

CCF [K] C C
Calculates and graphs cross correlations between two time series. If the number of lags is not specified, lags from $-(\sqrt{n} + 10)$ to K = $+(\sqrt{n} + 10)$, where n = the length of the series, are calculated.

LAGS K
Specifies the number of lags to display in the CCF plot. If you omit LAGS, Minitab uses a default number of lags. This is (K = $-(\sqrt{n} + 10)$ to K = $+(\sqrt{n} + 10)$), where K is the number of lags and n = the length of the series.

SCCF C
SCCF stores the cross correlations.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

GSAVE *"filename"*
Saves a plot of lags by cross correlations as the text string "filename".

DECOMP: Session command for performing decomposition on a time series

DECOMP C K

Performs classical decomposition on a time series, using either a multiplicative or an additive model. Classical decomposition decomposes a time series into trend, seasonal, and error components.

With DECOMP, list the column that contains the series, followed by a constant that indicates the seasonal length. If you do not specify a model, then Minitab uses the multiplicative model.

Decomposition displays a summary table and a set of plots. The summary table includes the trend equation, the seasonal indices, and three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD.

Decomposition generates three sets of plots:

- A time series plot that shows the original data, the fitted trend line, the trend plus seasonal fitted values, and forecasts
- Component analysis: A set of plots that illustrate how adjusting the series for different components affects the results
- Seasonal analysis: A set of plots that illustrate how the seasonal pattern affects the data

Model

ADDITIVE

By default, Minitab uses the multiplicative model. Use the multiplicative model when the size of the seasonal pattern in the data depends on the level of the data. This model assumes that as the data increase, so does the seasonal pattern. Most time series exhibit such a pattern.

The multiplicative model is: y_t = Trend * Seasonal + Error

The additive model is: y_t = Trend + Seasonal + Error

NOTREND

Excludes a trend component from the decomposition. You might want to do this if you have already detrended your data with TREND on page 962. Warning: If the data contain a trend component but you omit it from the decomposition, this can influence the estimates of the seasonal indices.

FIRST K

Specifies the seasonal period of the first data value. Enter a number to specify a starting value. For example, if you have monthly data and the first observation is in June, then enter 6 to set the seasonal period correctly. By default, the starting value is 1 because Minitab assumes that the first data value in the series corresponds to the first seasonal period.

Forecasts

FORECASTS K

Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot.

ORIGIN K

Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52. If you specify FORECASTS but no ORIGIN, Minitab generates forecasts from the end of the data. Minitab uses data up to the ORIGIN for fitting the trend model used to generate forecasts.

X-axis scale

INDEX or I

Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K

Use INCREMENT to increment the time scale.

STAMP C...C

Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month

Calendar unit subcommand	Range	Default START value
QUARTER or Q	1 to 4	Current quarter
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

TRENDS C
TRENDS stores the trend component data. This is the trend component; it does not contain the error or seasonal component.

DETREND C
DETREND stores the detrended data (has seasonal and error components, no trend).

SEASONALS C
SEASONALS stores the seasonal component data (has seasonal component, no error or trend).

DESEASONALS C
DESEASONALS stores the data which has had the seasonal component removed (has trend and error components, no seasonal).

RESIDUALS C
RESIDUALS stores the residuals in a column. Also called detrended data, you can use this data to do further analysis such as generating diagnostic plots using ACF on page 932. The detrended data equal the original data minus the fits.

FITS C
FITS stores the fits in a column. These are the values used to plot the trend line.

FSTORE C
FSTORE stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

Graphs

NOPLOT
Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM
GHISTOGRAM displays a histogram of the residuals.

GNORMAL
GNORMAL displays a normal probability plot of the residuals.

GFITS
GFITS plots the residuals versus the fitted values.

GORDER
GORDER plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
GFOURPACK displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
GVARIABLE displays a separate graph for the residuals versus each specified column.

Results

BRIEF K

Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

DES: Session command for performing double exponential smoothing

DES C

Performs Holt double exponential smoothing for a time series.

Use double exponential smoothing to smooth out noise and to forecast data that exhibit a trend. Both the Holt and Brown methods calculate a level component and a trend component at each period. Holt's method uses two smoothing parameters for updating these components. Double exponential smoothing uses the level and trend components to generate forecasts. Initial values for these components are obtained either by backcasting (if you omit the WEIGHT subcommand), or from a linear regression on time (if you use the WEIGHT subcommand).

With DES, list the column containing the time series. The time series cannot include any missing values. If you have missing data, see Missing values in exponential smoothing on page 1254.

The method that Minitab uses to calculate level and trend components depends on the inclusion of the WEIGHT subcommand. See for details.

Default output consists of a time series plot which displays the data and one-period-ahead forecasts, or fitted values. It also displays the smoothing constants used for level and trend, and three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD. Alternatively, you can generate a time series plot that displays the data and the smoothed values. Optional output includes a table containing the original series, the smoothed data, the fitted data and the errors.

Weight

WEIGHTS K

If you omit the WEIGHT subcommand, Minitab uses optimal weights, or smoothing parameters, which are computed by minimizing the sum of squared errors. It does this by using the equivalent ARIMA (0,2,2) model.

How you specify weights controls whether DES uses Holt's or Brown's method. Minitab uses Holt's method unless you specify the appropriate weights to use Brown's method. (Brown's method is equivalent to an equal-root ARIMA model. If you omit WEIGHTS, Minitab computes optimal weights, which will virtually never be the equivalent of an equal-root ARIMA model. This is why omitting WEIGHTS results in the Holt method.)

Without WEIGHTS, the trend and level components are initialized by backcasting. With WEIGHTS, the trend and level components are initialized using a linear regression on time.

The first weight updates the level component. Theoretically, this can be any number between 0 and 2, but it is normally a value between 0 and 1. The second weight updates the trend component. This can be any value between 0 and 100, although it is normally a value between 0 and 1.

Theoretically, the second weight can be any number greater than 0 and less than ((4/first weight) – 2) to preserve invertibility. But, because the latter value can be quite high, Minitab imposes a maximum value of 100 to ensure a more reasonable model. If you omit the WEIGHT subcommand, Minitab uses optimal weights, or smoothing parameters, which are computed by minimizing the sum of squared errors. It does this by using the equivalent ARIMA (0,2,2) model.

Forecasts

FORECASTS K

Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot with 95% prediction interval bands.

ORIGIN K

Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52 based on the moving average at period 48. If you omit the ORIGIN subcommand, Minitab generates forecasts from the end of the data.

X-axis scale

INDEX or I

Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K

Use INCREMENT to increment the time scale.

STAMP C...C

Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month
QUARTER or Q	1 to 4	Current quarter
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

SMOOTHED c

SMOOTHED stores the smoothed data.

LEVEL C

LEVEL stores the level components. The level component at time T equals the smoothed value at time T, while adding the level and trend components at time T equals the one-period-ahead forecast for time T + 1.

TREND C

TREND stores the trend components.

FITS C

FITS stores the fitted values. These are the one-period-ahead forecasts.

RESIDUALS C

RESIDUALS stores the residuals. The residual at time T is the difference between the actual data at time T and the fitted value at time T. These residuals are used to calculate MAPE, MAD, and MSD. If you store the residuals you can generate diagnostic plots using ACF on page 932.

FSTORE C

FSTORE stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

UPPER C

UPPER stores the upper 95% prediction limits for the forecasts.

LOWER C

LOWER stores the lower 95% prediction limits for the forecasts.

Graphs

SMPLOT

By default, Minitab plots predicted values and actual data. SMPLOT says to plot the smoothed values (the moving averages) and actual data.

NOPLOT
Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Results

BRIEF K
Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

DIFFERENCES: Session command for calculating differences

DIFFERENCES [K] C C
Calculates the differences between the elements of a column. Minitab subtracts from each row the element K rows above, where K is the lag specified, and stores the differences in a new column. The first K rows of the new column will contain the missing value symbol, *.

If C1 contains Z1, Z2, ..., Zn, then DIFFERENCES K C1 C2 puts asterisks (*) into rows 1 through K of C2, and puts Zi – Zi −k into row i, K + 1 < i < n. If K is omitted, then K = 1 is used.

LAG: Session command for calculating the lags of a column

LAG [K] C C

Calculates the lags of a column and stores them in a new column.

Moves the row elements of a column down K rows, where K is the lag specified, storing the result in a new column of the same length. There will be K missing value symbols, *, at the top of the output column. The output column has the same number of rows as the input column, so the last K values from the input column are not lagged. If K is omitted, then K = 1 is used.

If C1 contains Z1, Z2, ..., Zn, then `LAG K C1 C2` puts asterisks (*) into rows 1 through K of C2, and puts $Z_i - Z_{i-k}$ into row i, K + 1 < i < n.

MLAG: Session command to calculate lags of one or more columns

MLAG

Calculates the lags of one or more columns. Specify either the `AFTER` subcommand or the `NWORKSHEET` subcommand to specify to store the lags at the end of the active worksheet or in a new worksheet.

ROOT C

Specify a root column. Use multiple instances of the `ROOT` subcommand in the same instance of `MLAG` to calculate the lags for different root columns. Use the `LAGS` subsubcommand after each instance of the `ROOT` subcommand.

- `LAGS K or K ... K or K:K`: Specify one or more unique values that are shorter than the length of the root column that you specify with the `ROOT` subcommand. You can also use a colon to denote consecutive numbers. For example, if you enter *5:10*, the lag values are `5 6 7 8 9 10`. Each lag value results in a new column that contains the lags for that value. The name of the column is the name of the root column with "_LagK" appended, where K is the lag of the new column. For example, if the root column's name is "Root," and the lag value is 1, the lag column's name is "Root_Lag1." Each lag column moves the row elements of the root column down K rows. The top of the lag column has K missing value symbols (*). The lag column has the same number of rows as the original column, so the lag column does not contain the last K values from the original column.

Use one of the following subcommands to specify how to store the lags. Each instance of `MLAG` requires one of the following subcommands.

AFTER

Store the generated lag columns after the existing columns in the active worksheet.

NWORKSHEET

Store the generated lag columns in a new worksheet.

- `INROOT`: Store the root column in the new worksheet before the lag columns.

Results

TMETHOD

Displays the Method table, which includes a table of the root columns and lags, where each row corresponds to a column that is an argument of a ROOT subcommand.

MA: Session command for calculating a moving average

MA C K

Use moving averages to smooth out the noise in a time series and forecast future values of the series.

Minitab calculates a moving average by averaging consecutive groups of observations in a series. For example, suppose a series begins with the numbers 4, 5, 8, 9, and 10 and you use a moving average length of 3. The first two values of the moving average are missing. The third value of the moving average is the average of 4, 5, and 8; the fourth value is the average of 5, 8, and 9; the fifth value is the average of 8, 9, and 10; and so on.

You can use the linear moving averages method with a time series that exhibits a trend, as well as with moving average schemes involving more than two moving averages. First, compute and store the moving average of the original series. Then use the AVERAGES subcommand to compute and store the moving average of the previously stored column to obtain a second moving average.

You can also perform naive forecasting. In naive forecasting, the forecast for time T+1 is the data value at time T. There are two ways to do naive forecasting with Minitab:

- Execute MA and specify a moving average length of 1.
- Execute SES on page 957 and specify a moving average length of 1.

With MA, list the column containing the series, followed by a constant indicating the length of the moving average.

Default output consists of a time series plot which displays the data and one-period-ahead forecasts, or fitted values. Alternatively, you can generate a time series plot that displays the data and the smoothed values.

The uncentered moving average at time T is the one-period-ahead forecast, or fitted value, for time T+1. If you center the moving averages, these are adjusted accordingly when producing the fitted values.

Minitab also displays three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD. Optional output includes a table containing the series, moving average values, and the fits and residuals. If you generate forecasts, they are also listed, with corresponding lower and upper 95% prediction limits.

Centering

CENTER

If you use this subcommand, Minitab places the moving average values at the period which is in the center of the range rather than the end of it. This is called centering the moving average, and is done to position the moving average values at their central positions in time.

Forecasts

FORECASTS K

Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot with 95% prediction interval bands.

ORIGIN K

Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52 based on the moving average at period 48. If you omit the ORIGIN subcommand, Minitab generates forecasts from the end of the data.

X-axis scale

INDEX or I

Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K

Use INCREMENT to increment the time scale.

STAMP C...C

Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month
QUARTER or Q	1 to 4	Current quarter
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

SMOOTHED C
Stores the smoothed data.

FITS C
Stores the fitted values. The uncentered moving average at time T is the fitted value for time T + 1.

RESIDUALS C
Stores the residuals. The residual at time T is the difference between the actual data at time T and the fitted value at time T. These residuals are used to calculate MAPE, MAD, and MSD. If you store the residuals you can generate diagnostic plots using ACF on page 932.

FSTORE C
Stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

UPPER C
Stores the upper 95% prediction limits for the forecasts.

LOWER C
Stores the lower 95% prediction limits for the forecasts.

Graphs

SMPLOT
By default, Minitab plots predicted values and actual data. SMPLOT says to plot the smoothed values (the moving averages) and actual data.

NOPLOT
Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Results

BRIEF K

Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

PACF: Session command for calculating partial autocorrelation

PACF

Calculates and plots the partial autocorrelations of a time series.

PACF has several subcommands and generates a high-resolution partial autocorrelation function, including confidence limits for the correlations, in a graph. You can store the partial autocorrelations and associated t-statistics for up to the first 75 lags.

If you do not specify the number of lags with LAGS, Minitab uses a default number of lags:

- n / 4 for a series with less than or equal to 240 observations
- $\sqrt{n}$ + 45 for a series with more than 240 observations, where n is the number of observations in the series

LAGS K

Specifies the number of lags to display in the PACF plot. If you omit LAGS, Minitab uses a default number of lags. This is n / 4 for a series with less than or equal to 240 observations or $\sqrt{n}$ + 45 for a series with more than 240 observations, where n is the number of observations in the series.

SPACF C

Stores the partial autocorrelations.

TSTATS C

Stores the t-statistics.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

SES: Session command for performing single exponential smoothing

SES C

Use single exponential smoothing to smooth out the noise in a time series and forecast future values of the series.

This method uses an exponentially weighted average of all past values of the series to calculate the smoothed value at each period. Note: An exponentially weighted average, with weight a, is approximately equivalent to an unweighted moving average of length (2 – a)/a. If you do not specify a weight, Minitab computes an optimal weight, or smoothing parameter, based on your data (using WEIGHT). In this case, Minitab computes the initial smoothed value by backcasting. If you do specify a weight, Minitab uses the average of the first six observations for the initial smoothed value.

With SES, list the column containing the time series. The time series cannot include any missing values. If you have missing data, go to Missing values in exponential smoothing on page 1254.

You can also perform naive forecasting with SES. In naive forecasting, the forecast for time T+1 is the data value at time T. There are two ways to do naive forecasting with Minitab:

- Execute MA and specify a moving average length of 1.
- Execute SES on page 957 and specify a moving average length of 1.

Default output consists of a time series plot which displays the data and one-period-ahead forecasts, or fitted values. It also displays the smoothing constant (weight) used and three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD. Alternatively, you can generate a time series plot that displays the data and the smoothed values. Optional output includes a table containing the series, smoothed data, the one-period-ahead forecasts (or fitted values), and the one-period-ahead errors.

Weight

WEIGHT K

Single exponential smoothing uses an exponentially weighted average of all past values of the series to calculate the smoothed value at each period. If you omit the WEIGHT subcommand, Minitab computes and uses an optimal weight by minimizing the sum of squared errors. If you omit the WEIGHT subcommand, the initial smoothed value is found by backcasting.

Use WEIGHT to enter a specific weight. This must be a number between 0 and 2, although it is unusual to use a number greater than 1. By default, the initial smoothed value is the average of the first six observations. Use INITIAL to change the default. For more information, go to Computing weights, or smoothed values on page 1221.

INITIAL K

By default, if you use WEIGHT, Minitab uses the average of the first six observations for the initial smoothed value. Use INITIAL to specify the initial smoothed value.

Forecasts

FORECASTS K

Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot with 95% prediction interval bands.

ORIGIN K

Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52 based on the moving average at period 48. If you omit the ORIGIN subcommand, Minitab generates forecasts from the end of the data.

X-axis scale

INDEX or I

Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K

Use INCREMENT to increment the time scale.

STAMP C...C

Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month
QUARTER or Q	1 to 4	Current quarter

Calendar unit subcommand	Range	Default START value
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

SMOOTHED c

SMOOTHED stores the smoothed data. The smoothed value at time T is the fitted value for time T + 1.

FITS C

FITS stores the fitted values. The uncentered moving average at time T is the fitted value for time T + 1.

RESIDUALS C

RESIDUALS stores the residuals. The residual at time T is the difference between the actual data at time T and the fitted value at time T. These residuals are used to calculate MAPE, MAD, and MSD. If you store the residuals you can generate diagnostic plots using ACF on page 932.

FSTORE C

FSTORE stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

UPPER C

UPPER stores the upper 95% prediction limits for the forecasts.

LOWER C

LOWER stores the lower 95% prediction limits for the forecasts.

Graphs

SMPLOT

By default, Minitab plots predicted values and actual data. SMPLOT says to plot the smoothed values (the moving averages) and actual data.

NOPLOT

Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM

GHISTOGRAM displays a histogram of the residuals.

GNORMAL

GNORMAL displays a normal probability plot of the residuals.

GFITS

GFITS plots the residuals versus the fitted values.

GORDER

GORDER plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK

GFOURPACK displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C

GVARIABLE displays a separate graph for the residuals versus each specified column.

Results

BRIEF K

Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

TBOXCOX: Session command for performing a Box-Cox transformation on time series data

TBOXCOX C

TBOXCOX transforms a time series to try to make the variance of the series stationary. Stationary variance is a requirement for an ARIMA model. Use a time series plot to determine if the variance of a time series is stationary. If the time series has a pattern in the spread of the points, then the variance is not stationary.

PERIOD K

Specifies the seasonal length in the time series. For example, if you have an annual cycle and monthly data, enter 12 for K. K must be a positive integer within [2, floor(N/2)] where N is the row count of the column for `TBOXCOX`, omitting any missing values from the beginning or end of the data. The default value if you do not specify this subcommand is 2.

LULIMITS K K

Specifies the lower and upper limits for the search for the optimal lambda. The first K is the lower limit and must be less than the second K. Both values of K must be within [-5, 5]. This subcommand is not valid with `LAMBDA`. The default values if you do not specify `LULIMITS` or `LAMBDA` are -1, 2.

LAMBDA K

Specifies the value of λ for the Box-Cox transformation. K must be within [-5, 5]. This subcommand is not valid with `LULIMITS`.

TMETHOD

Displays the method table.

GTSPLOT [C]

Displays time series plots of the original and transformed time series data. Use the optional column argument to specify a column of labels for the x-axis of the graph. The optional column must be the same length as the time series data. The subcommand allows missing values in the optional column.

STORE C

Stores the transformed data in C.

TREND: Session command for performing a trend analysis

TREND C

Use trend analysis to fit a particular type of trend line to a time series or to detrend a time series.

You may fit one of four models: linear, quadratic, growth curve, and S-curve (Pearl-Reed logistic). The default is a linear trend model.

TREND generates a time series plot that shows the original data, the fitted trend line, and forecasts. Minitab also displays the fitted trend equation and three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD.

Model type

Use to fit a quadratic, exponential growth curve or S-curve model. If you do not enter one of these subcommands, Minitab uses the default model, the linear trend model. Use care when interpreting the coefficients from the different models, because they have different meanings.

QUADRATIC

Specifies the quadratic model.

GROWTH

Specifies the growth curve model.

SCURVE

Specifies the S-curve trend model.

Options

You can use these subcommands to incorporate knowledge learned from fitting the same trend model to prior data to obtain an "improved" fit to the present data. For example, suppose you fit a trend model to data covering each annual financial cycle. Further, suppose that the type of model chosen for this year is the same type chosen for last year. In that case, you may want to include the information from last year's model with this year's data.

This is how it works. First, fit last year's data with TREND using the appropriate model. Note the coefficients listed in the fitted trend equation displayed in the output. Next, fit this year's data with %TREND using the same model. Include the PRIORS subcommand, giving the coefficients you noted earlier as arguments.

Minitab generates the default time series plot, based only on this year's data, plus a second time series plot that shows trend lines for three models. Minitab displays the coefficients and accuracy measures for all three models.

This is how to interpret the three trend lines displayed in the second time series plot:

- Trend line 1 is based on this year's data only and is the same as the trend line in the first time series plot. Minitab calls this the "new line".
- Trend line 2 is based on the coefficients listed with PRIORS. The resulting fits generate the second trend line. Minitab calls this the "priors line".

- Trend line 3 is based on coefficients computed from a weighted average of the priors coefficients and the new coefficients. By default, these coefficients are computed by multiplying 0.2 times the new coefficient plus 0.8 times the prior coefficient. Minitab generates a third set of fits using these "smoothed" coefficients. Minitab calls this the "smoothed line". This smoothed trend line incorporates prior information and new information in much the same way that exponential smoothing works. In a sense, this smoothing of coefficients filters out some of the noise from the model parameters estimated in successive cycles.

Use WEIGHTS to change how the smoothed coefficients are computed. By default, Minitab multiplies each new coefficient by 0.2. To change this default, enter WEIGHTS with the value you want for each coefficient in the model.

PRIORS K K [K]
Specifies prior parameter values.

WEIGHTS K K [K]
Specifies weights of coefficients of current data for the weighted average fit.

Forecasts

FORECASTS K
Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot.

ORIGIN K
Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52. If you specify FORECASTS but no ORIGIN, Minitab generates forecasts from the end of the data. Minitab uses data up to the ORIGIN for fitting the trend model used to generate forecasts.

X-axis scale

INDEX or I
Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]
Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K
Use INCREMENT to increment the time scale.

STAMP C...C
Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month
QUARTER or Q	1 to 4	Current quarter
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0

Time unit subcommand	Range	Default START value
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

RESIDUALS C

RESIDUALS stores the residuals in a column. Also called detrended data, you can use this data to do further analysis such as generating diagnostic plots using ACF on page 932. The detrended data equal the original data minus the fits.

FITS C

FITS stores the fits in a column. These are the values used to plot the trend line.

FSTORE C

FSTORE stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

Graphs

NOPLOT

Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM

Displays a histogram of the residuals.

GNORMAL

Displays a normal probability plot of the residuals.

GFITS

Plots the residuals versus the fitted values.

GORDER

Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK

Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C

Displays a separate graph for the residuals versus each specified column.

Results

BRIEF K

Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

TSWINT: Session command for performing Holt-Winters seasonal exponential smoothing

TSWINT C K

Performs Holt-Winters seasonal exponential smoothing, using either a multiplicative or an additive model, to smooth and forecast data which exhibit both a trend and a seasonal pattern. Minitab uses the multiplicative model by default. Use the ADDITIVE subcommand to use an additive model.

This method employs a level component, a trend component, and a seasonal component at each period. It uses three weights, or smoothing parameters, to update the components at each period. Initial values for the level and trend components are obtained from a linear regression on time. Initial values for the seasonal component are obtained from a dummy-variable regression using detrended data.

With either command list the column containing the time series followed by a constant indicating the length of the seasonal pattern. This constant must be a positive integer greater than or equal to 2. The time series cannot include any missing values.

Default output consists of a time series plot which displays the data and one-period-ahead forecasts, or fitted values. It also displays the smoothing constants used for the level, trend and seasonal components, and three measures to help you determine the accuracy of the fitted values: MAPE, MAD, and MSD. Alternatively, you can generate a time series plot that displays the data and the smoothed values. Optional output includes a table containing the original series, the smoothed data, the fitted values and the errors.

Model

ADDITIVE

By default, Minitab uses the multiplicative model. Use the ADDITIVE subcommand with TSWINT to use an additive model.

Weights

WEIGHTS K K K

Specifies the weights for the level, trend, and seasonal components. By default, all three weights, or smoothing parameters, are set to 0.2. Since an equivalent ARIMA model exists only for a very restricted form of the Holt-Winters model, Minitab does not compute optimal parameters for TSWINT as it does for SES on page 957 and DES on page 946.

Use WEIGHTS to specify your own smoothing parameters. Each weight must be a number from 0 to 1.

Forecasts

FORECASTS K

Use FORECASTS with an integer to indicate how many forecasts you want. Forecasts appear in red on the time series plot with 95% prediction interval bands.

ORIGIN K

Use ORIGIN with a positive integer to specify a starting point for the forecasts. For example, if you specify 4 forecasts and 48 for the origin, Minitab computes forecasts for periods 49, 50, 51, and 52 based on the moving average at period 48. If you omit the ORIGIN subcommand, Minitab generates forecasts from the end of the data.

X-axis scale

INDEX or I

Numbers the x-axis with a single scale from 1 to n by ones (where n is the number of observations in the column containing the time series). INDEX is the default time scale for the x-axis. The default start value for INDEX is 1, and integers are valid values.

START K [K...K]

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands, and they should match the order the subcommands were given.

If you do not use START, the time series graph uses the default start values for the time unit subcommands. You must specify all time unit subcommands before you use START.

INCREMENT K

Use INCREMENT to increment the time scale.

STAMP C...C

Use the STAMP subcommand to place a date/time stamp on the x-axis. The argument must be a date/time column. You can use at least one and at most three columns of equal length to stamp the x-axis. The values in the date/time column are stamped on the time series graph in the same format as they appear in the date/time column. When you use the STAMP subcommand, you cannot use the START and INCREMENT subcommands and any of the time unit subcommands.

The optional arguments to the time unit subcommands let you change the default interval in which you measure data. This allows you to do a time series plot with, for example, bimonthly data, or data collected four times per minute.

For example, the following commands create a graph that shows the depth of a reservoir. Measurements of the depth were collected on a bimonthly basis (in months 1, 3, 5, 7, 9, and 11) starting in January of 1990. The argument for MONTH uses abbreviated notation to specify every other month from months 1 to 11 inclusive.

```
TSPLOT Depth;
 MONTH 1:11/2;
 YEAR;
 START 1 1990.
```

Calendar values for the x-axis

Labels the x-axis with calendar values. MONTH and YEAR account for different length months and leap years. Specify all the calendar unit subcommands before using any other subcommands in the graph.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Calendar unit subcommand	Range	Default START value
DAY or D	1 to a maximum of 31, depending on the month	Current day
MONTH or M	1 to 12	Current month
QUARTER or Q	1 to 4	Current quarter
YEAR or Y	1000 to 4000	Current year

Adding an additional calendar unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY

DAY; MONTH

DAY; MONTH; YEAR

MONTH

MONTH; YEAR

MONTH; QUARTER

MONTH; QUARTER; YEAR

QUARTER

QUARTER; YEAR

YEAR

Time of day values for the x-axis

Labels the x-axis with time values.

The following table shows the time unit subcommands, the range of values they can take, and their default start values.

Time unit subcommand	Range	Default START value
SECOND or SS	0 to 59	0

Time unit subcommand	Range	Default START value
MINUTE or MM	0 to 59	0
HOUR or HH	0 to 23	0

Adding an additional time unit subcommand to the same graph adds additional x-axes to the graph. You can use up to three subcommands in the following combinations:

DAY; HOUR

DAY; HOUR; MINUTE

HOUR

HOUR; MINUTE

HOUR; MINUTE; SECOND

MINUTE

MINUTE; SECOND

SECOND

Note that you can use DAY in combination with SECOND, MINUTE, and HOUR.

Storage

SMOOTHED C

Stores the smoothed data in the specified column. The smoothed value for time T is equal to the seasonal estimate for time T – p (where p is the seasonal period) multiplied by (or added to for an additive model) the level estimate for time T – 1.

LEVEL C

Stores the level components in the specified column.

TREND C

Stores the trend components in the specified column.

SEASONAL C

Stores the seasonal components in the specified column.

RESIDUALS C

Stores the one-period-ahead forecast errors in the specified column. These residuals are used to calculate MAPE, MAD, and MSD. If you store the residuals you can generate diagnostic plots using ACF on page 932.

FITS C

Stores the one-period-ahead forecasts in the specified column. The fitted value for time T is equal to the sum of the level and trend estimates for time T – 1, multiplied by (or added to for an additive model) the seasonal estimate for time T – p (where p is the seasonal period).

FSTORE C

Stores the forecasts in a column. Use this subcommand only if you also used the FORECASTS subcommand to generate forecasts.

UPPER C

Stores the upper 95% prediction limits for the forecasts.

LOWER C
Stores the lower 95% prediction limits for the forecasts.

Graphs

SMPLOT
By default, Minitab plots predicted values and actual data. SMPLOT says to plot the smoothed values (the moving averages) and actual data.

NOPLOT
Suppresses the default time series plot. If you do not use NOPLOT, Minitab displays the time series plot.

GHISTOGRAM
Displays a histogram of the residuals.

GNORMAL
Displays a normal probability plot of the residuals.

GFITS
Plots the residuals versus the fitted values.

GORDER
Plots the residuals versus the order of the data. The row number for each data point is shown on the x-axis (for example, 1 2 3 4... n).

GFOURPACK
Displays a layout of a histogram of the residuals, a normal probability plot of the residuals, residuals vs fitted values, and residuals vs order of the data.

GVARIABLE C...C
Displays a separate graph for the residuals versus each specified column.

Results

BRIEF K
Controls the amount of output. BRIEF can also be used as a main command. It then effects all commands that are controlled by BRIEF.

Value of K	Output that Minitab displays
0	Minitab displays no output, but performs all specified storage.
1	Minitab displays the summary table.
2	Minitab displays the summary and results table.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

Tables

TABLE: Session command for creating one-way, two-way, and multi-way tables using categorical variables

TABLES C...C

Displays one-way, two-way, and multi-way tables using categorical variables in C...C. TABLES uses two types of variables, categorical and associated. All the columns specified with TABLES must be the same length.

Categorical variables determine the cells of the table and are listed with TABLES. Categorical columns may be numeric or text. When using four categorical columns, a two-way table of the first entry (row) and second entry (column) is displayed for all possible combination of levels of the third and the fourth entry.

Associated variables are listed in the second set of eleven commands and determine summary statistics such as means, standard deviation, etc for each cell.

If no subcommands are used, the table displays the count for each cell which is the number of times a combination of row and column category occurs. If subcommands are used, then the counts are not displayed unless you specify the COUNTS subcommand. To calculate and store additional statistics such as skewness and kurtosis, go to STATS on page 144.

The TABLE command was developed in collaboration with James R. Allen of Allen Data Systems, Cross Plains, Wisconsin.

FREQUENCIES C

Allows the user to enter multiple occurrences in a single row of the worksheet. Usually each row of the worksheet corresponds to one case. The column must be of same length as the column specified in the main command. The column elements should be integers. Any value less than one, or a missing value is treated as 0. In textbooks, frequencies are sometimes called weights or case weights.

Statistics for categorical variables

These subcommands are for categorical variables specified in the main command.

COUNTS

Displays the counts of the total number of observations in each cell.

ROWPERCENTS

Displays the percentage each cell represents of the total observations in the row.

COLPERCENTS

Displays the percentage each cell represents of the total observations in the column.

TOTPERCENTS

Displays the percentage each cell represents of all the observations in the table.

Statistics for associated variables

These subcommands calculate statistics for each associated variable for each cell and margins (the rows and columns of a table). The columns can be numeric or date/time and must be the same length as the columns specified in the main command. Date/time columns are treated as numeric for all subcommands except MINIMUM and MAXIMUM. When you use MINIMUM and MAXIMUM, date/time columns display values in date/time format.

MEANS C...C
Calculates the arithmetic mean.

MEDIANS C...C
Calculates the median.

SUMS C...C
Calculates the sum of all the observations.

MINIMUMS C...C
Displays the smallest value.

MAXIMUMS C...C
Displays the largest value.

STDEV C...C
Calculates the standard deviation.

STATS C...C
Displays the same result as using the three subcommands N, MEANS, and STDEV.

DATA C...C
Displays a list of all the data in each cell except margins. The columns can be numeric, date/time, or text.

N C...C
Displays the number of nonmissing values.

NMISS C...C
Displays the number of missing values.

PROPORTION K [K] C...C
Displays the proportion of cases equal to the specified value (or in the specified range, if two K's are given). For example, PROP 2 5 C1 C2 gives two numbers for each cell, the proportion of cases where C1 is in the range 2 and 5, and the proportion of cases where C2 is in the range 2 and 5. You must have numeric data, and there cannot be any missing values to use the PROPORTION subcommand.

Options

These subcommands are for categorical variables. Use only columns that appeared on the main command.

CMISSING C...C
Includes missing values in calculating marginal counts and percentages.

DMISSING C...C
Missing values are not included in calculating marginal counts and percentages. Row, column, and total percentages are not calculated for the category 'Missing'.

Output

NOALL

NOALL suppresses the printing of marginal statistics.

ALL C...C

ALL prints marginal statistics over all levels of the specified categorical variables. The variables listed on ALL must also be listed on the TABLE command.

LAYOUT K K

Displays compact tables for reports. LAYOUT takes 2 numeric constant arguments. Both constants must be nonnegative integers. The second cannot be greater than 2. The sum of the two constants must be less than or equal to the number of columns specified with the TABLE command.

For example, if you use LAYOUT K1 K2, then the level combinations of the first K1 categorical variables will define the rows of the table, and the level combinations of the next K2 categorical variables will define the columns of the table. Any remaining variables are used to define multiple table.

TALLY: Session command for displaying a one-way table for each column

TALLY C...C

Displays a one-way table for each column.

If no subcommands are used, the table displays the count for each category. If subcommands are used, then counts are not displayed unless you specify COUNTS or ALL subcommand.

The data columns can be numeric or text and can contain any values.

Statistics for categorical variables

These subcommands are for categorical variables specified in the main command.

COUNTS

Displays the number of times a category occurs (count).

PERCENTS

Displays what percentage of the whole is accounted for by each category.

CUMCOUNTS

Displays the cumulative count for each category.

CUMPERCENTS

Displays the cumulative percentage for each category.

ALL

Displays the same result as using the four commands, COUNTS, PERCENTS, CUMCOUNTS, and CUMPERCENTS.

Storage

STORE C...C

Stores output from TALLY. If you specify n columns with TALLY and request m (COUNTS, CUMCOUNTS, PERCENTS, and CUMPERCENTS) statistics, then the number of columns you require to specify with the STORE subcommand is n + m * n. No statistics are stored for the missing values.

The first column you list, contains the data categories, and the rest store the requested statistics in the order in which the subcommands are given. For example, the following commands store the data categories in C2, the percents in C3, and the counts in C4.

```
TALLY C1;
  PERCENT;
  COUNT;
  STORE C2 C3 C4.
```

You can use the stored information to produce frequency plots, cumulative frequency plots, relative frequency plots, and cumulative relative frequency for integer data. For example, if you had integer data in column C1, you could produce a frequency plot with the following commands.

```
TALLY C1;
  STORE C2 C3.
  PLOT C3*C2
```

TCHISQUARE: Session command for performing a chi-square goodness-of-fit test

TCHISQUARE

TCHISQUARE performs a chi-square goodness-of-fit test. If your data are the actual categories in a column, categorical data are in raw form, specify the column. If the column contains missing data, then that row is excluded from all calculations.

OBSERVED C

OBSERVED K...K

Use OBSERVED to specify the observed counts for each category. All values entered must be nonnegative. The number of observed counts must be greater than one. You cannot enter missing values.

LEVELS C

LEVELS K...K

Use LEVELS with OBSERVED to specify the levels of each category. You can specify the levels in a column or as constants. The number of values in a column or constants must equal the number of categories.

PROPORTIONS [C]

PROPORTIONS [K]...[K]

Specify the proportions to calculate the expected value for each category in a column or as constants. If you do not use this subcommand, Minitab assumes equal proportion across categories. Missing values are not allowed.

COUNTS C

COUNTS K...K
Specifies the historical counts to calculate the expected value for each category. You can enter historical counts as constants or in a column. Values must be nonnegative. Missing values are not allowed.

Graph

GBAR
Displays a bar chart of the observed and expected values.

GHISQ
Use GCHISQ to display a bar chart of each category's contribution to the chi-square value.

PARETO
Use PARETO with GCHISQ to display a bar chart that orders the bars from largest to smallest of each category's contribution to the chi-square value.

Results

RTABLE
Displays a table with observed and expected counts and the test results.

XTABS: Session command for displaying one-way, two-way, and multi-way tables for categorical variables

XTABS [C...C]
Displays one-way, two-way, and multi-way tables. Categorical columns may be numeric or text, or date/time, and may contain any values. All columns must be the same length. When a categorical column is text, value order of the column is used.

UNSTACKED C...C
Specifies that the data are arranged as a table of summarized count data in C...C.

RLEVELS C
RLEVELS allows you to specify a column of names for the categories of the row variable.

CNAME K
CNAME allows you to specify a name for the column variable.

FREQUENCIES C
Allows you to enter multiple occurrences in a single row of the worksheet. Usually each row of the worksheet corresponds to one case. The column must be of same length as the column specified in the main command. The column elements should be integers. Any value less than one, or a missing value is treated as 0. In textbooks, frequencies are sometimes called weights or case weights.

Categorical statistics

Each subcommand displays a statistics for categorical variables listed with XTABS.

COUNTS
Displays the count of the total number of observations (missing and nonmissing) in each cell.

ROWPERCENTS
Displays the percentage each cell and margin represents of the total observations in the row.

COLPERCENTS
Displays the percentage each cell and margin represents of the total observations in the column.

TOTPERCENTS
Displays the percentage each cell and margin represents of all the observations in the table.

Chi-square analysis

CHISQUARE
Calculates likelihood ratio and Pearson's chi-square tests of independence between the rows and columns of a two-way table. The output includes the test statistics, the number of degrees of freedom, the p-value of the tests and the number of cells with expected count less than 5 (if any) for each test. If expected cell count is less than one, no p-value is displayed. The subcommand works only when using LAYOUT 1,1. (Yates' correction for 2 x 2 tables is not used.)

EXPECTED
Displays the expected cell counts in the table.

RESIDUALS
Displays the residuals (Observed - Expected) in the table.

SRESIDUALS
Displays the standardized residuals in the table.

ARESIDUALS
Displays the adjusted residuals in the table.

XRESIDUALS
Displays each cell's contribution to the Pearson chi-square statistic.

Tests for 2x2 tables

FISHER
Displays Fisher's exact test p-value for each two-way table.

MCNEMAR
Displays McNemar's test for each two-way table. McNemar's procedure tests whether paired proportions are different. For example, you can use McNemar's test to determine whether a training program changes the proportion of participants who answer a question correctly.

The results include the estimated difference in proportions, a confidence interval for the difference, and the p-value for the difference.

MHCTEST
Displays the estimates of the common odds ratio and the Mantel-Haenszel-Cochran test statistics with corresponding p-value. This subcommand applies only when three or more variables are specified with the main command. There is only one test rather than one per table. The first two categorical variables must have 2 levels each.

Measure of agreement

KAPPA
Displays Cohen's Kappa statistics for rater reliability for each two-way table. This subcommand applies only when the output is in a two-way table format.

Measures of association

These subcommands are for calculating measures of association.

VSQUARE
Displays Cramer's V-squared statistics.

PRE
Displays Proportional Reduction of Error measures of association, Goodman and Kruskal lambda and tau for each two-way table. Two versions of each statistics is displayed, one for each choice of dependent and independent variables.

CONCORDANT
Displays the number of concordant pairs, discordant pairs, and p-values for Goodman and Kruskal's gamma, Somer's d, and Kendall's tau-b measures of association for each two-way table.

CORRELATION
Displays the Pearson's r and Spearman's rho rank correlation coefficients for each two-way table.

Options

Use only columns that appeared on the main command.

DMISSING C...C
Missing values are not included in calculating marginal counts and percents. Row, column, and total percents are not calculated for the category 'Missing'.

CMISSING C...C
Includes missing values in calculating marginal counts, percents, expected counts, residuals, and chi-square tests.

Output

NOALL
NOALL suppresses the printing of marginal statistics.

ALL C...C
ALL prints marginal statistics over all levels of the specified categorical variables. The variables listed on ALL must also be listed on the TABLE command.

LAYOUT K K

Use LAYOUT to print compact tables for reports. LAYOUT takes 2 numeric constant arguments. Both constants must be nonnegative integers. The second cannot be greater than 2. The sum of the two constants must be less than or equal to the number of columns specified on the TABLE command.

For example, LAYOUT K1 K2, the first K1 categorical variables will define the rows of the table, while the level combinations of the next K2 categorical variables will define the columns of the table. Any remaining variables are used to define multiple table.

Nonparametric Analysis

1-Sample Sign

STEST: Session command for performing a 1-sample sign test

STEST [K] C...C

Performs a 1-sample sign test separately for each column. STEST tests the null hypothesis that the population median is equal to the hypothesized value (K). If you do not specify a hypothesized median in K, STEST compares the sample median to 0.

Use SINTERVAL on page 979 to calculate a sign confidence interval for the median.

ALTERNATIVE K

Specifies a one-sided alternative hypothesis for the test. K takes two values, 1 and −1. If K = −1, then H1: M < hypothesized value. If K = +1, then H1: M > hypothesized value. If this command is not used, Minitab performs a two-sided test such that H1: M ≠ hypothesized value.

SINTERVAL: Session command for calculating a sign confidence interval

SINTERVAL [K] C...C

Calculates a sign confidence interval, separately for each column. You can specify a confidence level on the command line. For example, if you enter the following command, Minitab calculates 90% confidence intervals. If you do not specify a confidence level, then SINTERVAL calculates a 95% confidence interval.

```
SINTERVAL 0.90 C1
```

Use STEST on page 979 to perform a sign test for the median.

1-Sample Wilcoxon

WTEST: Session command for performing a 1-sample Wilcoxon test

WTEST [K] C...C

Performs a 1-sample Wilcoxon signed-rank test of the median separately for each column. If you do not specify a hypothesized median, WTEST compares the sample median to 0.

Use WINTERVAL on page 980 to calculate a confidence interval for the median.

ALTERNATIVE K

Specifies a one-sided alternative hypothesis for the test. K takes two values, +1 and −1. If K = −1, then H1: M < hypothesized value. If K = +1, then H1: M > hypothesized value. If this command is not used, Minitab performs a two-sided test such that H1: M ≠ hypothesized value.

WINTERVAL: Session command for calculating a Wilcoxon confidence interval

WINTERVAL [K] C...C

WINTERVAL calculates Wilcoxon confidence interval separately for each column.

You can specify a confidence level on the command line. For example, if you enter the command, Minitab calculates a 90% confidence interval.

```
WINTERVAL 0.90 C1
```

If you do not specify a confidence level, WINTERVAL gives a 95% confidence interval.

Use WTEST on page 979 to perform a 1-sample Wilcoxon test for the median.

MANN-WHITNEY: Session command for performing a Mann-Whitney test

MANN-WHITNEY [K] C C

MANN-WHITNEY tests for differences between two population medians.

The two columns must be numeric and can be unequal in length. You can specify a confidence level on the command line. The confidence level must be between 0.5 – 1. If you do not specify a confidence level, a default level of 95% is used.

```
MANN 0.90 C1 C2
```

ALTERNATIVE K

Specifies a one-sided alternative hypothesis for the test. K takes two values, +1 and −1. If K = −1, then H_1: $M_1 < M_2$, where M_1 and M_2 are the medians for population 1 and 2 respectively. If K = +1, then H_1: $M_1 > M_2$. If this command is not used, Minitab performs a two-sided test such that H_1: $M_1 \neq M_2$.

FRIEDMAN: Session command for performing a Friedman test

FRIEDMAN C C C [C[C]]

Performs a Friedman test with of a randomized block experiment, in the stated order, response in C, treatment in C, blocks in C, and optionally, you can store residuals in C and fits in C. The null hypothesis for the analysis is that the treatment has no effect.

The treatment and blocks columns can be numeric or text, and can contain any values. You must have exactly one observation per treatment-block. Missing data are not allowed.

KRUSKAL-WALLIS: Session command for performing a Kruskal-Wallis test

KRUSKAL-WALLIS C C

Performs a Kruskal-Wallis test with response in C and factors in C.

Performs a nonparametric analysis of the one-way analysis of variance. KRUSKAL-WALLIS tests the null hypothesis that there are no differences among the k populations.

The factor column can be numeric or text, and can contain any value. The factors do not need to be in any special order.

MOOD: Session command for performing a Mood's median test

MOOD C C [C[C]]

Performs a Mood's median test with response in C, factors in C, and optionally, you can store residuals in C and fits in C.

Performs a nonparametric analysis of a one-way layout. MOOD tests the null hypothesis that the population medians are equal.

The factor levels can be numeric or text. The levels do not need to be in any special order.

RUNS: Session command for performing a runs test

RUNS [K] C...C

Performs a run test separately on each column. RUNS tests the null hypothesis that the data come from a random process by using the comparison criterion K. If K is not specified, the mean of the scores is used.

You can have missing data (*) at the beginning or end of the data column or columns, but not in the middle.

WALSH: Session command for calculating pairwise averages

WALSH C C [C C]

Calculates Walsh averages for data in C and stores in a column C. Optionally, you can store the two indices which defines the Walsh averages in columns C and C.

Calculates the Walsh average of all possible pairs of values, including each value with itself.

Let $x_1, x_2, \ldots, x_n$ be the observations. The Walsh average, $(x_i + x_j) / 2$, has indices i and j. If you specify columns to store the indices, the value of i is put in the first column and j in the second column. If you have n observations, there will be $n (n + 1) / 2$ Walsh averages.

This command is useful for nonparametric tests and confidence intervals.

WDIFF: Session command for calculating pairwise differences

WDIFF C C C [C C]

Calculates differences between all the pairs of elements from two columns C and C and store in a column C. Optionally, you can store the two indices which defines the pairwise differences in columns C and C.

Calculates all possible differences between pairs by subtracting a value in the second column from the corresponding value in the first column.

Let $x_1, x_2, \ldots, x_n$ be the values in the first column, and $Y_1, Y_2, \ldots, Y_m$ be the values in the second. WDIFF calculates all the differences, $(x_i – y_j)$. If you specify columns to store the indices, the value of i is put in the first column and j in the second column.

These differences are useful for nonparametric tests and confidence intervals. For example, the point estimate given by MANN-WHITNEY on page 980 can be computed as the median of the differences.

WSLOPE: Session command for calculating pairwise slopes

WSLOPE C C C [C C]

Calculates slope between pairs of points in columns C (x-coordinate)and C (y-coordinate) and stores in a column C. Optionally, you can store the two indices which defines the pair in columns C and C.

WSLOPE calculates the slope between every pair of points and stores the slopes in the column listed on the command line. Each row of the y-x columns defines a point in the plane. If you specify columns to store the indices, the two row numbers used to calculate the slope are put in the corresponding row of these two columns.

If any observations are missing or if the slope cannot be defined (for example, the slope of a line parallel to the y-axis), then the slope is computed as missing. If there are n rows in the input columns, there will be n (n – 1) / 2 slopes in the output column.

Equivalence Test

TOST: Session command for performing an equivalence test

You can use an equivalence test to determine whether the means for product measurements or process measurements are close enough to be considered equivalent.

Small differences between products are not always functionally or practically important. For example, a difference of 1 mg in a 200 mg dose of a drug is unlikely to have any practical effect. When you use an equivalence test, you specify how large the difference must be to be considered important. Smaller differences are considered unimportant. Minitab then tests two separate null hypotheses:

- The difference is less than or equal to your lower limit for equivalence.
- The difference is greater than or equal to your upper limit for equivalence.

If both null hypotheses are rejected, then the difference falls within your equivalence interval and you can claim that the means for the products are equivalent.

This information is also conveyed graphically in an equivalence plot. If the confidence interval for the difference is completely within your equivalence limits, then you can claim equivalence.

Minitab provides equivalence tests for the following experimental designs.

1-sample equivalence test
Compare the mean of a product or a process to a standard.

2-sample equivalence test
Compare two product means or two process means using 2 independent samples.

Equivalence test with paired data
Compare two product means or two process means when observations are paired, such as when measurements for two treatments are taken from the same person or part.

Equivalence test for 2x2 crossover design
Compare treatment means for data that was collected in a 2x2 crossover study. 2x2 crossover studies are commonly used to test whether individual participants experience equivalent effects on 2 different drugs.

1-sample equivalence test

`TOST C K`
Performs a one-sample equivalence test to determine whether the test mean (test sample in C) is equivalent to the target value K.

`TOST K K K K`
Performs a one-sample equivalence test on summarized data, where each K specifies the following (in the following order):

- Sample size
- Sample mean
- Sample standard deviation

- Target value

2-sample equivalence test

TOST C C

Performs a 2-sample equivalence test to determine whether the test mean is equivalent to the reference mean; the first column C contains the test sample, and the second column C contains the reference sample.

TOST C C

Performs a 2-sample equivalence test with the following data layout:

- Sample data are in the first column C.
- Sample IDs (with exactly 2 unique values) are in the second column C.

STACKED

Specifies that the data are stacked.

[REFERENCE K}

The optional constant, REFERENCE K, specifies which sample ID value represents the reference sample.

TOST K K K K K K

Performs a 2-sample equivalence test with summarized data, where each K specifies the following (in the following order):

- Test sample: Sample size
- Test sample: Mean
- Test sample: Standard deviation
- Reference sample: Sample size
- Reference sample: Mean
- Reference sample: Standard deviation

Paired equivalence test

TOST C C

Performs an equivalence test for paired data with test responses in the first column C and paired reference responses in the second column C.

PAIRED

2x2 crossover design

TOST C C C C K

Performs an equivalence test on data from a 2x2 crossover design with data in separate columns as follows (in the following order):

- Response values for period 1 of sequence 1
- Response values for period 2 of sequence 1
- Response values for period 1 of sequence 2

- Response values for period 2 of sequence 2

Value of K	Order
1	Specifies that the treatment order for sequence 1 is reference followed by test
2	Specifies that the order for sequence 1 is test followed by reference

TOST C C C K

Performs an equivalence test on data from a 2x2 crossover design with sequence ID in the first column C, data for period 1 in the second column C, paired data for period 2 in the third column C, and the order of the treatments for sequence 1 specified by K.

Value of K	Order
1	Reference followed by test
2	Test followed by reference

TLABEL K K

Specifies alternative labels for the reference and test treatments (for example, "Standard" and "Experimental") as the first K and the second K, respectively (the default labels are "Reference" and "Test").

Equivalence limits

LLIMIT and ULIMIT specify the equivalence limits. If you specify only one limit (LLIMIT or ULIMIT), then Minitab performs a one-sided test. If you specify RATIO or LOG, then both equivalence limits must be greater than 0.

LLIMIT K

LLIMIT specifies the lower equivalence limit as numeric value K.

ULIMIT K

ULIMIT specifies the upper equivalence limit as numeric value K.

RMEAN

RMEAN specifies that the limit represents a proportion of the reference mean or target value. For example, check to specify the limit as 0.2 multiplied by the mean of the reference sample rather than as the fixed value of 0.2. RMEAN cannot be used with LOG or RATIO. The reference mean or target value must be greater than 0.

Hypotheses

These subcommands specify the hypothesis for 2-sample, paired, and 2x2 crossover designs.

By default, Minitab tests whether the difference, test mean – reference mean (or target value), is within the equivalence limits.

RATIO

RATIO specifies to test whether the ratio, test mean / reference mean, is within the equivalence limits. For RATIO, both equivalence limits must be greater than 0 and the reference mean cannot be equal to 0.

LOG

LOG specifies to test the ratio by testing the difference between the means after transforming the data by taking the natural log of each data value. (All data values and equivalence limits must be greater than 0.)

Options

UEVARIANCE
Specifies that equal variances are not assumed when performing a two-sample equivalence test.

ALPHA K
Specifies the significance level for the test (also called alpha or α).

CIALT
Displays the (1 – 2α)*100 % confidence interval instead of the default (1 – α)*100 % confidence interval.

Graphs

GINTERVAL
Displays the equivalence plot which includes a confidence interval for the difference or ratio.

GHISTOGRAM
For 1-sample and 2-sample designs, GHISTOGRAM displays a histogram of each sample. For paired data, GHISTOGRAM displays a histogram of the paired differences.

GINDPLOT
For 1-sample and 2-sample designs, GINDPLOT displays an individual value plot that shows the data from each sample. For paired data, GINDPLOT displays an individual value plot of the paired differences.

GBOXPLOT
For 1-sample and 2-sample designs, GBOXPLOT displays a boxplot that shows the data from each sample. For paired data, GBOXPLOT displays a boxplot of the paired differences.

GPROFILE
Displays a subject profile plot for paired designs and 2x2 crossover designs. Symbols for related responses are connected by lines to help reveal response patterns.

GPERIOD
Displays a scatterplot of period 1 responses versus period 2 responses for 2x2 crossover designs.

GMEAN
Plots the mean response for each period of each sequence in a 2x2 crossover design.

Results

TMETHOD
Displays the method table, which shows the details of the analysis, such as which sample is the test sample and which is the reference sample.

TDSTATISTICS
Displays descriptive statistics, which includes the sample size, the mean, and the standard deviation for each sample.

TEFFECTS
Displays the effects table for 2x2 crossover designs. Minitab tests for a carryover effect, a treatment effect, and a period effect.

TDMEAN

If you choose a difference hypothesis, then TDMEAN displays the difference between the means and the confidence interval for that difference. If you choose a ratio hypothesis, then TDMEAN displays the ratio of the means and the confidence interval for that ratio.

TTEST

Displays the results of the hypothesis tests.

Power and Sample Size Analysis

FDESIGN: Session subcommand for power and sample size for a general full factorial design

FDESIGN K K

Use FDESIGN with the POWER command to calculate power, the number of replicates, or the maximum value of the difference between main effect means for a general full factorial design.

NLEVELS K...K

Specifies the number of levels for each factors in the model as K. K is an integer greater than or equal to 2.

TORDER K

Includes terms in the model up to order K.

FITB

Specifies that blocks is a term in the model.

REPS K...K

Specifies one or more numbers of replicates.

You must specify two of REPS, MAXDIFFERENCE, and POWER. Minitab calculates the third.

MAXDIFFERENCE K...K

Specifies the maximum difference between main effect means.

You must specify two of REPS, MAXDIFFERENCE, and POWER. Minitab calculates the third.

POWER K...K

Specifies one or more power values.

You must specify two of REPS, MAXDIFFERENCE, and POWER. Minitab calculates the third.

ALPHA K

Specifies the significance level of the test. The default is 0.05.

SIGMA K

Specifies an estimate for the standard deviation.

Storage

These subcommands store information about the power and sample size calculations.

SREPLICATES C

Stores the number of replicates.

SMAXDIFFERENCE C

Stores the maximum difference between main effect means.

SPOWER C
Stores the power values.

STOTALRUNS C
Stores the total number of runs.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus maximum difference for the specified test conditions.

NSIZE K...K
NSIZE generates additional power curves based on specified number of replicates.

FFDESIGN: Session subcommand for power and sample size for a 2-level factorial design

FFDESIGN
Use with the POWER command to calculate power, number of replicates, effect size, or number of center points for a two-level full or fractional factorial design.

REPS K...K
Specifies one or more numbers of replicates.

You must specify two of REPS, MAXDIFFERENCE, and POWER. Minitab calculates the third.

EFFECT K...K
Specifies one or more alternative hypotheses as a difference in means or effect size.

POWER K...K
Specifies one or more power values.

You must specify two of REPS, MAXDIFFERENCE, and POWER. Minitab calculates the third.

CPBLOCK K...K
Specifies the number of center points per block. The default is 0.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

BLOCKS K
Specifies the number of blocks. The default is 1.

OMIT K
Specifies the number of terms that you are going to omit from the default model.

FITC
Specifies that, if there are center points, there will be a term for center points in the model.

FITB
Specifies that, if there are blocks, there will be a term for blocks in the model.

Storage

The following subcommands store information about the power and sample size calculations.

SREPLICATES C
Stores the number of replicates.

SEFFECT C
Stores the effect sizes.

SPOWER C
Stores the power values.

SCPBLOCK C
Stores the number of center points per block.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus effect for the specified test conditions.

The following optional subsubcommands generate additional power curves on one graph to compare experimental designs.

NSIZE K...K
Generates additional power curves based on specified number of replicates.

NCTRPT K...K
Generates additional power curves based on specified number of center points per block.

ONERATE: Session subcommand for power and sample size for a 1-sample Poisson rate test

ONERATE
Use with the POWER command to calculate power, sample size, or comparison rate for a one-sample Poisson rate procedure.

RNULL K
Specifies the hypothesized null rate. K must be greater than 0.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

LENGTH K
Specifies the length of the observation period as K. The default is 1.

Power, sample size, and comparison rate

You must specify values for any two of the following subcommands. Minitab solves for the third.

POWER K...K
Specifies one or more power values.

SAMPLE K...K
Specifies one or more sample sizes.

RCOMPARE K...K
Specifies one or more comparison rates. K must be greater than 0.

Storage

SSAMPLE C
Stores the sample size.

SRCOMPARE C
Stores comparison rates.

SPOWER C
Stores power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison rate for the specified test conditions.

NSIZE K...K
Generates additional power curves on one graph to compare sample sizes.

ONEVARIANCE: Session command for power and sample size for a 1 variance test

ONEVARIANCE

Use with the POWER command to calculate power, sample size, or significant ratio for a 1 variance test.

STDEV

Indicates that input and output values refer to standard deviation instead of variance, where applicable.

ALPHA K

Specifies the significance level of the test. The default is 0.05.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Power, sample size, and ratio

You must specify values for any two of the following subcommands. Minitab solves for the third.

POWER K...K

Specifies one or more power values.

SAMPLE K...K

Specifies one or more sample sizes.

RATIO K...K

Specifies one or more alternative hypotheses in terms of the ratio between the comparison variability and the hypothesized variability.

Storage

SSAMPLE C

Stores the sample size.

SRATIO C

Stores the ratios.

SPOWER C

Stores power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison rate for the specified test conditions.

NSIZE K...K
Generates additional power curves on one graph to compare sample sizes.

ONEWAY: Session command for power and sample size for one-way ANOVA

ONEWAY
Use with the POWER command to calculate power, sample size, or maximum difference between means for a one-way analysis of variance procedure with K levels.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

Power, sample size, and maximum difference

You must specify values for any two of the following subcommands. Minitab solves for the third.

POWER K...K
Specifies one or more power values.

SAMPLE K...K
Specifies one or more sample sizes.

MAXDIFFERENCE K...K
Specifies the maximum difference between means.

Storage

SSAMPLE C
Stores the sample size.

SPOWER C
Stores power values.

SMAXDIFFERENCE C
Stores maximum differences.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison rate for the specified test conditions.

NSIZE K...K
Generates additional power curves on one graph to compare sample sizes.

PBDESIGN: Session subcommand for power and sample size for Plackett-Burman design

PBDESIGN K K
Use with the POWER command to calculate power, number of replicates, effect size, or number of center points for a Plackett-Burman design with K factor levels and K corner points.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

OMIT K
Specifies the number of main effect terms to omit from the full main effects model.

FITC
Specifies that, if there are center points, there will be a term for center points in the model.

Power, replicates, effect, and center points

You must specify values for any two of the following subcommands. Minitab solves for the third.

POWER K...K
Specifies one or more power values.

REPS K...K
Specifies one or more numbers of replicates.

EFFECT K...K
Specifies one or more alternative hypotheses as a difference in means or effect size.

CPBLOCK K...K
Specifies the number of center points. The default is 0.

Storage

SREPLICATES C
Stores the number of replicates.

SEFFECT C
Stores the effect sizes.

SPOWER C
Stores one or more power values.

SCPBLOCK C
Stores the number of center points.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison rate for the specified test conditions.

NSIZE K...K
Generates additional power curves on one graph to compare sample sizes.

NCTRPT K...K
Generates additional power curves based on the specified number of center points.

PONE: Session subcommand for power and sample size for a 1 proportion test

PONE
Use with the POWER command to calculate power, sample size, or comparison proportion for a 1 proportion test.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

PNULL K
Specifies the value of the proportion under the null hypothesis.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Comparison proportions, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

PCOMPARE K...K
Specifies one or more comparison proportions.

SAMPLE K...K
Specifies one or more sample sizes.

POWER K...K
Specifies one or more power values.

Storage

SSAMPLE C
Stores the sample sizes.

SPCOMPARE C
Stores the comparison proportions.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

POWER: Session command for power and sample size

POWER
Calculates power or other experimental parameters such as sample size. Each of the following subcommands contains several subsubcommands. Click on the corresponding link for a listing of those subsubcommands.

Use Minitab's power and sample size capabilities to perform both prospective and retrospective power analysis.

Prospective power analysis
Use a prospective study before collecting data to assess the power of an experimental design while you can do something about it. For example, you can increase the power of a design by increasing the sample size or by taking measures to decrease the error variance.

Retrospective power analysis
Use a retrospective study after collecting data to help understand the power of the tests that you have performed. For example, suppose you conduct an experiment and the data analysis does not reveal any statistically significant results. You can calculate power based on the minimum difference (effect) you wish to detect. If the power to detect this difference is low, you may want modify your experimental design to

increase the power and continue to evaluate the same problem. However, if the power is high, you may want to conclude that there is no meaningful difference (effect) and discontinue further experimentation.

ZONE

Calculates power, sample size, or significant difference for a 1-sample Z-test. For more information, go to ZONE on page 1010.

TONE

Calculates power, sample size, or significant difference for a 1-sample t-test. For more information, go to TONE on page 1003.

TTWO

Calculates power, sample size, or significant difference for a 2-sample t-test. For more information, go to TTWO on page 1006.

TPAIRED

Calculates power, sample size, or significant difference for a paired t-test. For more information, go to TPAIRED on page 1005.

PONE

Calculates power, sample size, or alternative proportion for a 1-sample proportion test. For more information, go to PONE on page 996.

PTWO

Calculates power, sample size, or significant difference for a 2-sample proportion test. For more information, go to PTWO on page 999.

ONERATE

Calculates power, sample size, or significant difference for a 1-sample Poisson rate test. For more information, go to ONERATE on page 991.

TWORATE

Calculates power, sample size, or significant difference for a 2-sample Poisson rate test. For more information, go to TWORATE on page 1008.

ONEVARIANCE

Calculates power, sample size, or significant ratio for a one variance test. For more information, go to ONEVARIANCE on page 993.

TWOVARIANCE

Calculates power, sample size, or significant ratio for a two variances test. For more information, go to TWOVARIANCE on page 1009.

TOST

Calculates power, sample size, or difference (or ratio) for an equivalence test. For more information, go to TOST on page 1004.

ONEWAY

Calculates power, sample size, or maximum difference between means for a one-way analysis of variance (ANOVA) test. For more information, go to ONEWAY on page 994.

FFDESIGN

Calculates power, number of replicates, effect size, or number of center points for two-level full and fractional factorial designs. For more information, go to FFDESIGN on page 990.

PBDESIGN

Calculates power, number of replicates, effect size, or number of center points for a Plackett-Burman design. For more information, go to PBDESIGN on page 995.

FDESIGN

Calculates power, number of replicates, or maximum value of difference between main effect mean for a general full factorial design. For more information, go to FDESIGN on page 989.

PTWO: Session subcommand for power and sample size for a 2 proportion test

PTWO

Use with the POWER command to calculate power, sample size, or comparison proportion for a 2 proportion procedure.

ALPHA K

Specifies the significance level of the test. The default is 0.05.

PBASELINE K

Specifies the baseline proportion as K.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Comparison proportions, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

PCOMPARE K...K

Specifies the comparison proportion.

SAMPLE K...K

Specifies one or more sample sizes. Sample sizes are per group.

POWER K...K

Specifies one or more power values.

Storage

SSAMPLE C
Stores the sample sizes.

SPCOMPARE C
Stores the comparison proportions.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

SSCI: Session command for estimating sample size

SSCI
Estimates the sample size needed to achieve a specific margin of error for confidence intervals for the mean, standard deviation, variance, proportion, Poisson rate, or Poisson mean. Conversely, you can estimate the margin of error based on sample size.

NMEAN
Use NMEAN to estimate the margin of error or sample size based on a confidence interval from a normal distribution.

SIGMA K [K]
Use SIGMA K to specify the standard deviation planning value. Use SIGMA [K] to specify whether the population standard deviation is known.

Value of [K]	Population standard deviation
0	Population standard deviation is not known
1	Population standard deviation is known

NSTDEV K
Use NSTDEV to estimate the margin of error or the sample size with a standard deviation planning value K based on a confidence interval from a normal distribution.

NVARIANCE K
Use NVARIANCE K to estimate the margin of error or the sample size with variance planning value K based on a confidence interval from a normal distribution.

BPROPORTION K

Use BPROPORTION to estimate the margin of error or the sample size with proportion planning value K based on a confidence interval from a binomial distribution. K must be greater than 0 but less than 1.

PRATE K

Use PRATE to estimate the margin of error or the sample size with rate planning value K based on a confidence interval from a Poisson distribution.

LENGTH K

Use LENGTH to specify the length of the observation. K > 0.

PMEAN K

Use PMEAN to estimate the margin or error or the sample size based on a confidence interval from a Poisson distribution.

MERROR K...K

MERROR specifies that you want to estimate the sample size based on the margins of error K...K.

SAMPLE K...K

SAMPLE specifies that you want to estimate the margin of errors for the confidence interval based on sample sizes K...K.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

SESTIMATE C

Stores sample size or margin of error (depending on what you are estimating) in C, one row for each estimate.

SSTI: Session command for sample size for tolerance intervals

SSTI

Estimates the sample size or margin of error for exact parametric and nonparametric tolerance intervals.

Use to determine the sample size that is required to achieve a specific margin of error for a tolerance interval. Conversely, you can calculate the margin of error based on sample size.

MERROR K...K

Specifies K...K as the margins of error for which you want to calculate sample size. Each K must be less than 100 – PPERCENT K. For example, if K for PPERCENT is 95, then each K for MERROR must be less than 5.

SAMPLE K...K

Specifies K...K as the sample sizes for which you want to calculate margins of error.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

PPERCENT K

Specifies K as the minimum percentage of the population that you want the tolerance interval to cover. The default is 95.

MEPROB K

Specifies K as the margin of error probability. The following statistics define the precision of the tolerance interval:

MERROR

The margin of error measures the additional percentage of the population, beyond the desired PPERCENT, that might be included in the interval.

MEPROB

The margin of error probability is the probability that the actual interval content exceeds PPERCENT by MERROR or more.

Choose a value that is greater than 0 and less than or equal to 0.5. Common values for the margin of error probability include 0.01, 0.05, and 0.1.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Results

TMETHOD

Displays the method table which includes the confidence level, the percentage of the population in the interval, and the margin of error probability.

TRESULTS

Displays the results table which includes the sample sizes and associated margins of error. Minitab displays the results for both the normal and nonparametric methods. The results for the nonparametric method include the achieved confidence and the achieved error probability.

TONE: Session command for power and sample size for a 1-sample t-test

TONE

Use with the POWER command to calculate power, sample size, or significant difference for a 1-sample t-test. For a paired t procedure, proceed as if conducting a 1-sample t-test on the distribution of differences.

ALPHA K

Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)

Specifies an estimate for sigma.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Difference, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

DIFFERENCE K...K

Specifies one or more alternative hypotheses in terms of the difference from the null hypothesis mean.

SAMPLE K...K

Specifies one or more sample sizes.

POWER K...K

Specifies one or more power values.

Storage

SDIFFERENCE C

Stores the differences.

SSAMPLE C

Stores the sample sizes.

SPOWER C

Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

TOST: Session command for power and sample size for an equivalence test

TOST K
Use with the POWER command to calculate power, sample size, or difference (or ratio) for one of the following equivalence tests.

Value of K	Equivalence test design
0	Paired design
1	1-sample design
2	2-sample design
2, with the CROSSOVER subcommand	2x2 crossover design

CROSSOVER
Use for a 2x2 crossover design.

RATIO
By default, Minitab performs calculation for the test of the difference. To perform calculations for the test of the ratio by taking the log transformation of the data, use RATIO.

SIGMA K (required)
Specifies an estimate for sigma.

ALPHA K
Specifies the significance level (also called alpha or a) of the test. The default is 0.05.

LLIMIT K
LLIMIT specifies the lower equivalence limit as numeric value K. If RATIO is used, then both equivalence limits must be greater than 0.

ULIMIT K
ULIMIT specifies the upper equivalence limit as numeric value K. If RATIO is used, then both equivalence limits must be greater than 0.

Difference, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

DIFFERENCE K...K
Specifies one or more difference values. If the RATIO subsubcommand is also used, then use DIFFERENCE to specify the ratio values rather than difference values. All values for DIFFERENCE must be within the equivalence limits specified by LLIMIT and ULIMIT.

SAMPLE K...K
Specifies one or more sample sizes.

POWER K...K
Specifies one or more power values.

Storage

SDIFFERENCE C
Stores the differences.

SSAMPLE C
Stores the sample sizes.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

TPAIRED: Session subcommand for power and sample size for a paired t-test

TPAIRED
Use with the POWER command to calculate power, sample size, or significant difference for a paired t-test.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Difference, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

DIFFERENCE K...K
Specifies one or more alternative hypotheses in terms of the difference from the null hypothesis mean.

SAMPLE K...K
Specifies one or more sample sizes.

POWER K...K
Specifies one or more power values.

Storage

SDIFFERENCE C
Stores the differences.

SSAMPLE C
Stores the sample sizes.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

TTWO: Session subcommand for power and sample size for a 2-sample t-test

TTWO
Use with the POWER command to calculate power, sample size, or significant difference for a 2-sample t-test.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Difference, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

DIFFERENCE K...K
Specifies one or more alternative hypotheses expressed in terms of the difference between the means.

SAMPLE K...K
Specifies one or more sample sizes. Sample sizes are per group.

POWER K...K
Specifies one or more power values.

Storage

SDIFFERENCE C
Stores the differences.

SSAMPLE C
Stores the sample sizes.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

TWORATE: Session subcommand for power and sample size for a 2-sample Poisson rate test

TWORATE

Use with the POWER command to calculate power, sample size, or comparison rate for a 2-sample Poisson rate test.

ALPHA K

Specifies the significance level of the test. The default is 0.05.

ALTERNATIVE K

Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

RBASELINE K

Specifies the baseline rate as K. K must be greater than 0.

LENGTH K

Specifies the length of the observation period as K. The default is 1.

Comparison rate, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

RCOMPARE K...K

Specifies one or more comparison rates. K must be greater than 0.

SAMPLE K...K

Specifies one or more sample sizes.

POWER K...K

Specifies one or more power values.

Storage

SCOMPARE C

Stores the comparison rates.

SSAMPLE C

Stores the sample sizes.

SPOWER C

Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

TWOVARIANCE: Session subcommand for power and sample size for a 2 variances test

TWOVARIANCE
Use with the POWER command to calculate power, sample size, or significant ratio for a 2 variances test.

STDEV
Indicates that input and output values refer to standard deviation instead of variance, where applicable.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Type of test

Specifies which test to use to calculate the power and sample size results.

LEVENE
Use Levene's test when data are from a continuous, but not necessarily normal, distribution.

FTEST K...K
Use the F-test only when data are from a normal distribution.

BONETT K
You can also use Bonett's test when data are from a continuous, but not necessarily normal, distribution. However, the power calculations for Bonett's test are very sensitive to kurtosis, so you must enter a good estimate of the kurtosis value for your data. Enter the true common kurtosis of the parent populations as K. Do not enter the "excess kurtosis". (Excess kurtosis = kurtosis – 3.)

Power, sample size, and ratio

You must specify values for any two of the following subcommands. Minitab solves for the third.

POWER K...K
Specifies one or more power values.

SAMPLE K...K
Specifies one or more sample sizes.

RATIO K...K
Specifies the ratio between comparison and baseline variability.

Storage

SSAMPLE C
Stores the sample size.

SRATIO C
Stores the ratios.

SPOWER C
Stores power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison rate for the specified test conditions.

NSIZE K...K
Generates additional power curves on one graph to compare sample sizes.

ZONE: Session subcommand for power and sample size for a 1-sample Z-test

ZONE
Use with the POWER command to calculate power, sample size, or significant difference for a 1-sample Z-test.

ALPHA K
Specifies the significance level of the test. The default is 0.05.

SIGMA K (required)
Specifies an estimate for sigma.

ALTERNATIVE K
Enter K to specify the direction of the alternative hypothesis.

Value of K	Format of the alternative hypothesis
–1	Less than
0 (default)	Not equal to
1	Greater than

Difference, sample size, and power

You must specify values for any two of the following subcommands. Minitab solves for the third.

DIFFERENCE K...K
Specifies one or more alternative hypotheses in terms of the difference from the null hypothesis mean.

SAMPLE K...K
Specifies one or more sample sizes.

POWER K...K
Specifies one or more power values.

Storage

SDIFFERENCE C
Stores the differences.

SSAMPLE C
Stores the sample sizes.

SPOWER C
Stores the power values.

Graphs

GPCURVE
Displays power curves. Power curves plot every combination of power versus comparison proportion for the specified test conditions.

NSIZE K...K
Generates additional power curves based on specified sample sizes.

Graphs

CHART: Session command for creating a bar chart

CHART C...C

CHART *F*(C...C)

CHART *F*(C...C) * C

CHART (C...C) * C

Use to compare categories of data. Each bar can represent a count for a category, a function of a category, or summary values from a table.

The following commands create two bar charts: counts of flaws and counts of grade.

```
CHART 'Flaws' 'Grade'.
```

The following commands create two bar charts: the mean of mileage for each quality category and the mean of PSI for each quality category.

```
CHART MEAN('Mileage' 'PSI')*'Quality'.
```

The following commands create a bar chart of the amount of money lost for each flaw.

```
CHART 'Losses'*'Flaws';
 SUMMARIZED.
```

Chart

STACK

Specifies how to display categories when you have multiple grouping levels. K refers to the stack order based on a column's value order.

Value of K	Stack order
1	Stacks the first variable on the bottom and the remaining variables in ascending order.
2	Stacks the first group on the top and the remaining in groups descending order.

SUMMARIZED

Denotes that data are summarized. That is, one numeric or date/time value describes each category.

Scale

CUMULATIVE

Use a cumulative y-scale. Thus, the height of the third bar, for example, is the sum of the observations in the first three categories. The subcommand SLEVEL designates on which scale level both PERCENT and CUMULATIVE operate.

INCREASING and DECREASING are performed before CUMULATIVE and PERCENT.

INCREASING

Present the groups in the chart in increasing order based on the y-axis values.

INCREASING and DECREASING are performed before CUMULATIVE and PERCENT.

INCREASING and DECREASING do not affect the ordering within clusters, stacks, or cluster/stack combinations; they instead move entire clusters, stacks, or cluster/stack combinations into order by the sum of the values in each outermost scale level category.

DECREASING

Present the groups in the chart in decreasing order based on the y-axis values.

INCREASING and DECREASING are performed before CUMULATIVE and PERCENT.

INCREASING and DECREASING do not affect the ordering within clusters, stacks, or cluster/stack combinations; they instead move entire clusters, stacks, or cluster/stack combinations into order by the sum of the values in each outermost scale level category.

PERCENT

Use a percent scale on the y-axis so that all bars at the designated level total 100%. The subcommand SLEVEL designates within which scale level both PERCENT and CUMULATIVE operate.

INCREASING and DECREASING are performed before CUMULATIVE and PERCENT.

SLEVEL [K]

Designates the scale level within which both PERCENT and CUMULATIVE operate.

Value of K	Scale level
0 (default)	All the data
1	Within groups of the outermost scale (must have 2 or more levels)
2	Within groups of the second from the outermost scale level (must have 3 or more levels)
3	Within groups of the third from the outermost scale level (must have 4 levels)

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Data view

BAR: Session subcommand for representing data values with bars on page 1076

CONNECT: Session subcommand for connecting points with lines on page 1085

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

HMAP: Session command for creating a heatmap

HMAP [C...C]

Displays a heatmap for each column C. The column can be text or numeric and must be the same length. If the column is a text column, the only functions you can use are PEQUAL, N, and NMISSING. Missing values are ignored in calculations except when the function is N or NMISSING. If you do not specify any columns, Minitab uses the number of nonmissing rows to define the gradient.

You must specify a function and at least one of RVARIABLE or CVARIABLE.

RVARIABLE C...C

Specify the categorical variables that appear on the rows. The columns can be text or numeric and must be the same length as the column you specify with HMAP. You can enter up to five columns. The order of the levels of the categorical variables in the heatmap follow the value order. Missing values appear as another categorical level.

CVARIABLE C...C

Specify the categorical variables that appear on the columns. The columns can be text or numeric and must be the same length as the column you specify with HMAP. You can enter up to five columns. The order of the levels of the categorical variables in the heatmap follow the value order. Missing values appear as another categorical level.

Functions

MEAN

The heatmap uses means to summarize the contents of the cells. Requires a numeric column for HMAP.

SUM

The heatmap uses sums to summarize the contents of the cells. Requires a numeric column for HMAP.

MEDIAN

The heatmap uses medians to summarize the contents of the cells. Requires a numeric column for HMAP.

PERCENTILE K
The heatmap uses percentiles to summarize the contents of the cells. K must be a numeric value between 0 and 100. Requires a numeric column for HMAP.

STDEV
The heatmap uses standard deviations to summarize the contents of the cells. Requires a numeric column for HMAP.

MINIMUM
The heatmap uses minimums to summarize the contents of the cells. Requires a numeric column for HMAP.

MAXIMUM
The heatmap uses maximums to summarize the contents of the cells. Requires a numeric column for HMAP.

PBETWEEN K1 K2
The heatmap uses the percent of observations from K1 to K2 to summarize the contents of the cells. K1 and K2 must be numeric values where K1 is less than or equal to K2. The percentage includes the observation if the observation equals K1 or K2. Requires a numeric column for HMAP.

PEQUAL K...K
The heatmap uses the percent of observations equal to any of K....K to summarize the contents of the cells. K can be text or numeric, but the values you enter for K must be the same type of data as the column you specified with HMAP.

N
The heatmap uses the number of non-missing cells to summarize the contents of the cells.

NMISSING
The heatmap uses the number of missing cells to summarize the contents of the cells.

Options

GRADIENT
Select the color scale for the rectangles. By default, Minitab uses DIVERGING with no arguments if you do not specify GRADIENT.

DIVERGING[K]
Minitab colors rectangles with high values red and rectangles with low values blue. If you specify K, then the gradient is symmetric around K.

SEQUENTIAL[K]
If K = -1, Minitab colors rectangles with low values dark blue and rectangles with high values light blue and light gray. This option could be used to highlight rectangles with low defect rates or to minimize cost. If K = 1, Minitab colors rectangles with high values dark blue and rectangles with low values light blue and light gray. This option could be used to highlight rectangles with more productivity or to maximize revenue.

MULTIPLE C...C
Creates separate graphs for each group created by the grouping columns C...C. The grouping columns can be numeric or text, but must be the same length as the columns that you specify with HMAP. Minitab treats missing values as a separate group.

INDEPENDENT
Creates separate plots for each group created by the individual columns specified with MULTIPLE independent of the other columns. For example, suppose you specify two columns with MULTIPLE. The first column

contains the gender, either male or female, and the other contains the marriage status, either married or single. If you specify INDEPENDENT, Minitab creates a graph for each group created by the columns individually:

- Male
- Female
- Single
- Married

If you do not specify INDEPENDENT, Minitab creates a graph for each group created by all the columns:

- Male and Single
- Male and Married
- Female and Single
- Female and Married

SAME 3

When you specify multiple columns with HMAP, use this subcommand to use the same gradient scale for each heatmap created for the different columns.

WAFERPLOT: Session command for creating a wafer plot

WAFERPLOT C*C*C

Displays a wafer plot where the first C is the response variable, the second C is the y variable, and the third C is the x variable. The columns must be numeric and have at least 1 non-missing value. At least 1 row of the data must have all non-missing values. The columns must all be the same length. You must specify one function.

Functions

FMEAN

The wafer plot uses means to summarize the contents of the cells.

SUM

The wafer plot uses sums to summarize the contents of the cells.

FMEDIAN

The wafer plot uses medians to summarize the contents of the cells.

PERCENTILE K

The wafer plot uses percentiles to summarize the contents of the cells. K must be a numeric value between 0 and 100.

STDEV

The wafer plot uses standard deviations to summarize the contents of the cells.

FMINIMUM
The wafer plot uses minimums to summarize the contents of the cells.

FMAXIMUM
The wafer plot uses maximums to summarize the contents of the cells.

PBETWEEN K1 K2
The wafer plot uses the percent of observations from K1 to K2 to summarize the contents of the cells. K1 and K2 must be numeric values where K1 is less than or equal to K2. The percentage includes the observation if the observation equals K1 or K2.

PEQUAL K...K
The wafer plot uses the percent of observations equal to any of K....K to summarize the contents of the cells. K must be numeric.

N
The wafer plot uses the number of non-missing cells to summarize the contents of the cells.

NMISSING
The wafer plot uses the number of missing cells to summarize the contents of the cells.

Multiple graphs

PANEL C
Display multiple plots in one plot where C is the variable that separates the plots. C must not contain more than 25 distinct values and must be the same length as the columns you enter with WAFERPLOT. C cannot contain all missing values. You cannot specify both PANEL and SEPARATE.

SEPARATE C
Display each plot in a separate plot where C is the variable that separates the plots. C must not contain more than 25 distinct values and must be the same length as the columns you enter with WAFERPLOT. C cannot contain all missing values. You cannot specify both PANEL and SEPARATE.

Labels

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

Note NODSUBTITLE and NODFOOTNOTE are not available for wafer plot.

Data view

SYMBOL
Use with one of the following subcommands to specify the attributes for the cell borders. NOBORDER is the default if you do not specify SYMBOL.

NOBORDER
Suppresses the border on the individual cells.

ECOLOR K

Specify the color of the borders on the individual cells where K is a number between 0 and 129. For more information on which number corresponds to which color, go to Numbers for colors to use in session commands on page 1255.

Options

GRADIENT

Use with one of the following subcommands to select the color scale for the rectangles. The plot uses `DIVERGING` with no arguments unless you specify `GRADIENT`.

DIVERGING[K]

Minitab colors rectangles with high values red and rectangles with low values blue. If you specify K, then the gradient is symmetric around K.

SEQUENTIAL[K]

If K = -1, then low values are dark blue and high values are light blue and light gray. This option could be used to highlight rectangles with low defect rates or to minimize cost. If K = 1, Minitab colors rectangles with high values dark blue and rectangles with low values light blue and light gray. This option could be used to highlight rectangles with more productivity or to maximize revenue.

GMIN K and GMAX K

Specify the minimum and maximum value for the gradient scale with K. If you do not specify a minimum or maximum, then the gradient scale uses the full data range of the defect column.

Note If you specified `PBETWEEN` or `PEQUAL`, the wafer plot uses a percentage scale for the gradient. In these cases, the value that you enter for `GMIN` and `GMAX` should be between 0 and 1.

VIRIDIS, MAGMA, PLASMA, INFERNO, TURBO

Provide different color schemes for the gradient.



AXLABEL K [C]

AXLABEL K *["string"]*

Generates customized axis labels. K specifies the axis to be affected and its label. You can enter a string inside of double quotes to specify the label or you can enter a column with text.

Value of K	Axis
1	X-axis
2	Y-axis

ADISPLAY K

Specifies the side and display of the axis label.

Value of K	How the axis label is displayed
0	Hides the axis label
1	Displays the axis label on the low side
2	Displays the axis label on the high side

PLOT: Session command for creating a scatterplot

Scatterplots illustrate the relationship between two variables by plotting one against the other on the y- and x-axes. You can also use PLOT with a date/time variable to create a time series plot, including time series plots with irregular intervals (which cannot be handled by TSPLOT).

PLOT with no subcommands displays symbols. Use SYMBOL if you want symbols to be displayed with CONNECT, PROJECT, AREA, LOWESS, or REGRESS.

Command language	Scatterplots that are displayed
`PLOT Revenue*Year`	One graph of Revenue by Year
`PLOT Revenue*Year Expense* Year`	Two graphs: Revenue by Year and Expense by Year
`PLOT (Revenue Expense)*Year`	Two graphs: Revenue by Year and Expense by Year
`PLOT (Revenue Expense)*(Quarter Year)`	Four graphs: Revenue by Quarter, Revenue by Year, Expense by Quarter, and Expense by Year
`PLOT (Sales Advertis)*Year; OVERLAY.`	One graph of Sales and Advertis by Year with different symbols representing Sales and Advertis

PLOT C * C

Creates a scatterplot of y (1st column) by x (2nd column).

PLOT C * C ... C * C

Creates multiple scatterplots.

Scale

XYSAME

Specifies the same scale for the x- and y-axes.

If XYSAME is used, the MIN and MAX subcommands to SCALE on page 1149 are ignored.

XYMIN K

Use the optional XYMIN subcommand to specify the minimum for the scales.

XYMAX K

Use the optional XYMAX subcommand to specify maximum for the scales.

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

CONNECT: Session subcommand for connecting points with lines on page 1085

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot on page 1122

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

REGRESS: Session subcommand for fitting a regression line to data on a graph on page 1148

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

NPLOT: Session command for creating a binned scatterplot

Use to investigate the relationship between a pair of continuous variables when the data set contains many observations.

NPLOT C*C C*C...C*C

Displays a scatterplot for each pair of numeric columns. The columns for each pair of columns must have the same number of rows.

BINNED

Specify a binned scatterplot.

MEAN C

Specify a column that defines the color of a bin by the mean of C for the observations in the bin. If you do not specify `MEAN`, Minitab defines the color of a bin by the number of observations in the bin.

Options

SCALE K

Specify either the x scale (K=1) or the y scale (K=2).

LOGSK

Perform a log transformation of the scale specified with SCALE. Enter a value in K to specify the base of the logarithm used to transform the scale. You must enter a numeric value between 2 and 10. A logarithmic scale linearizes logarithmic relationships by changing the axis, so that the same distance represents different changes in value across the scale.

MULTIPLE C...C

Creates separate plots for each group created by the grouping columns C...C. The grouping columns can be numeric or text, but must be the same length as the columns that you specify with NPLOT. Minitab treats missing values as a separate group.

INDEPENDENT

Creates separate plots for each group created by the individual columns specified with MULTIPLE independent of the other columns. For example, suppose you specify two columns with MULTIPLE. The first column contains the gender, either male or female, and the other contains the marriage status, either married or single. If you specify INDEPENDENT, Minitab creates a plot for each group created by the columns individually:

- Male
- Female
- Single
- Married

If you do not specify INDEPENDENT, Minitab creates a plot for each group created by all the columns:

- Male and Single
- Male and Married
- Female and Single
- Female and Married

GRADIENT

Select the color scale for the bins. By default, Minitab uses DIVERGING with no arguments if you do not specify GRADIENT. You can only specify GRADIENT and its subcommands when you also specify BINNED.

DIVERGING[K]

Minitab colors bins with high values red and bins with low values blue. If you specify K then the gradient is symmetric around K.

SEQUENTIAL[K]

If K = -1 Minitab colors bins with low values dark blue and bins with high values light blue and light gray. This option could be used to highlight bins with low defect rates or to minimize cost. If K = 1 Minitab colors bins with high values dark blue and bins with low values light blue and light gray. This option could be used to highlight bins with more productivity or to maximize revenue.

SAME K [K] [K]

Use the same scale on each graph. Specify multiple arguments to set more than one scale the same.

- To use the same x-scale, set K = 1.
- To use the same y-scale, set K = 2.
- To use the same color gradient scale on a binned scatterplot for different pairs of x y-variables, set K = 3. For the same pair of x y-variables, the gradient scale will always be the same.

GCORRELATION: Session command for creating a correlogram

GCORRELATION C...C

Displays a correlogram for columns C...C. You must enter at least two columns and all columns must have the same number of rows. The columns must be numeric. Minitab calculates the Pearson correlation coefficient for each pair of columns.

Options

MULTIPLE C...C

Creates separate plots for each group created by the grouping columns C...C. The grouping columns can be numeric or text, but must be the same length as the columns that you specify with GCORRELATION. Minitab treats missing values as a separate group.

INDEPENDENT

Creates separate plots for each group created by the individual columns specified with MULTIPLE independent of the other columns. For example, suppose you specify two columns with MULTIPLE. The first column contains the gender, either male or female, and the other contains the marriage status, either married or single. If you specify INDEPENDENT, Minitab creates a graph for each group created by the columns individually

- Male
- Female
- Single
- Married

If you do not specify INDEPENDENT, Minitab creates a graph for each group created by all the columns:

- Male and Single
- Male and Married
- Female and Single
- Female and Married

GRADIENT

Select the color scale for the rectangles. If you do not specify GRADIENT, the gradient type depends on the calculation of the Pearson correlation coefficients.

- If the coefficients are positive and negative, the gradient type is DIVERGING.
- If the coefficients are all positive, the gradient type is SEQUENTIAL 1.
- If the coefficients are all negative, the gradient type is SEQUENTIAL -1.

DIVERGING

Minitab colors rectangles with high values red and rectangles with low values blue. The gradient is symmetric around 0.

SEQUENTIAL[K]

If K = -1, Minitab colors rectangles with low values dark blue and rectangles with high values light blue and light gray. If K = 1, Minitab colors rectangles with high values dark blue and rectangles with low values light blue and light gray.

MATRIXPLOT: Session command for creating a matrix of plots

A matrix plot is a two-dimensional array of individual plots that provides an efficient way to see meaningful relationships among a number of variables at once.

MATRIXPLOT C...C

Displays a matrix of plots. The list of variables defines both the row (y) variables and column (x) variables in the plot matrix. You can use 2 to 20 numeric or date/time columns of equal length.

The following command language displays a 3x3 matrix plot.

```
MATRIXPLOT 'Weight' 'Head.L' 'Head.W'.
```

MATRIXPLOT (C...C) * (C...C)

Displays a graph of each y versus each x (draftsman plot or casement display). The variables before the * define the row (y) variables, and the variables after the * define the column (x) variables of the matrix. You can have 1–20 numeric or date/time columns of equal length for both y and x.

The following command language displays a matrix plot of each y (Pulse1 and Pulse2) versus each x (Height, Weight, and Activity).

```
MATRIXPLOT ('Pulse1' 'Pulse2')*('Height' 'Weight' 'Activity').
```

Matrix labels

The following subcommands specify where to display labels, and apply only to the matrix of plots form of MATRIXPLOT.

DIAGONAL (default)

Labels the variables in the diagonal.

BOUNDARY

Labels the variables on the outside of the frame.

Graphs

Specifies the graphs to display.

FULL (default)

Displays all graphs.

LL

Displays graphs in the lower left triangle.

UR

Displays graphs in the upper right triangle. The default sides for tick and axis lines and labels shift to the upper and right sides, respectively.

Alternating sides for tick marks

NOALTERNATE

Turn off the default method of alternating sides for tick display.

By default, Minitab places the ticks of adjacent panels on alternating sides of the graph. Invoking NOALTERNATE places ticks on the same side of adjacent panels.

Margin size

MARGIN K

Specify the margin size separating all panels.

By default, margins for a matrix plot are 0. You can enter a value from 0 to 0.25. The maximum valid value depends on how many panels there are. If you use a margin argument that is too large, Minitab will use the largest possible value.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

CONNECT: Session subcommand for connecting points with lines on page 1085

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot on page 1122

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

REGRESS: Session subcommand for fitting a regression line to data on a graph on page 1148

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

BUBBLEPLOT: Session command for creating a bubble plot

BUBBLEPLOT C * C

Use a bubble plot to explore the relationships among three variables on a single plot. Like a scatterplot, a bubble plot plots a y-variable versus an x-variable. However, the symbols (also called bubbles) on the bubble plot vary in size. The area of each bubble represents the value of a third variable.

You need three columns of equal length, in numeric or date/time format. You can also use up to three columns of categorical data for grouping and up to two columns of categorical data for paneling.

Minitab does not plot an observation in the following cases:

- An x-value, y-value, or bubble size value is missing.
- The value in BSIZE C is less than or equal to zero.

BSIZE C

Specifies the size of the bubbles.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

BSYMBOL [C...C]

Specifies up to three columns of categorical variables for grouping. Minitab displays all groups on the same graph.

COLOR K...K

COLOR C

Specifies the color for the bubble groups, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each group or store the symbol colors in C.

BSCALE K

Specifies to make K the maximum bubble size. The bubble that represents the largest value in column BSIZE C is displayed at size K. The area of the remaining bubbles is scaled proportionally. K must be greater than 0 and less than or equal to 100. The default is 20.

NOSCALE

Specifies to interpret the values in BSIZE C as Minitab symbol sizes. The maximum Minitab symbol size is 100.

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

Multiple graphs

PANEL: Session subcommand for paneling graphs on page 1134

Data options

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

MARGPLOT: Session command for creating a marginal plot

MARGPLOT C * C

Displays a marginal plot of y * x.

Use to display a scatterplot with separate marginal plots of the x- and y-variables. You can choose to have histograms, boxplots, or dotplots as the marginal plots.

The data must be two numeric or date/time columns of equal length.

MARGIN K

Specifies the variables for which to create marginal plots. If you do not use MARGIN, both x and y are used (the default).

Value of K	Variables that the plot is created for
1	x-variables only
2	y-variables only

OPTION K

Specifies the type of marginal plot to display:

Value of K	Type of marginal plot
1 (default)	Histograms
2	Boxplots
3	Dotplots

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

HLABEL: Session subcommand for labeling histogram bars on a marginal plot with y-axis values on page 1103

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

BAR: Session subcommand for representing data values with bars on page 1076

DOT: Session subcommand for displaying a symbol for each data value on page 1094

IQRBOX: Session subcommand for displaying an interquartile range box on a boxplot on page 1114

OUTLIER: Session subcommand for displaying a symbol for each outlier on a boxplot on page 1133

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graphs, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

HISTOGRAM: Session command for creating a histogram

HISTOGRAM C...C

Displays a separate histogram for each column.

HISTOGRAM divides sample values into many intervals called bins. Bars represent the number of observations falling within each bin (its frequency). Observations that fall exactly on an interval boundary are included in the interval to the right (or left, if the last bin).

HISTOGRAM with no subcommands displays bars. Use BAR if you want bars to be displayed with SYMBOL, PROJECT, AREA, or DISTRIBUTION.

Y-scale type

CUMULATIVE

Specifies a cumulative scale for the y-axis. When CUMULATIVE is used with a frequency scale, for example, the height of the third bar represents the total number of observations in the first three intervals. You can also use CUMULATIVE with PERCENT to get a cumulative percent scale. CUMULATIVE may not be used with DENSITY.

DENSITY

Specifies a density scale for the y-axis in which the area of a bar is equal to the proportion of observations that fall within that bin. The height of each bar is calculated as:

$$\frac{\text{number of observations in interval}}{(\text{number of observations in all intervals}) * (\text{width of interval})}$$

A density scale is useful for viewing probability density functions.

PERCENT

Specifies a percent scale for the y-axis in which the height of a bar represents the percentage of observations that fall within that bin.

YFREQUENCY

Specifies a frequency scale for the y-axis in which the height of a bar represents the number of observations that fall within that bin.

If you use more than one of DENSITY, PERCENT, and YFREQUENCY for the same graph, only the last one is used.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

CUTPOINT, MIDPOINT, and NINTERVAL: Session subcommands for specifying cutpoints and midpoints on page 1087

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

BAR: Session subcommand for representing data values with bars on page 1076

DISTRIBUTION: Session command for fitting a distribution on page 1091 You must use HPARAM with DISTRIBUTION.

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot on page 1122

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

DOTPLOT: Session command for creating a dotplot

Use to assess and compare distributions by plotting the values along a number line.

The x-axis for a dotplot is divided into many small intervals, or bins. Data values falling within each bin are represented by dots.

If possible, Minitab displays a dot for each observation. If there is insufficient space, a dot represents multiple observations with a footnote indicating the maximum number of observations represented by the dots.

The data must be numeric or date/time. The categorical grouping data can be numeric, date/time, or text.

DOTPLOT C...C

Displays a separate graph for each C.

The following command language displays two dotplots: one of Sales and one of Advertis.

```
DOTPLOT 'Sales' 'Advertis'.
```

DOTPLOT (C...C) * C

Displays a separate graph for each C on the left, with a dotplot for each category of the C on the right.

The following command language displays a dotplot of yearly sales.

```
DOTPLOT 'Sales'*'Year'.
```

Dotplot options

STACK K

Specifies how to display categories when you have multiple grouping levels. K refers to the stack order based on a column's value order.

Value of K	How categories are displayed
1 (default)	Stacks the first variable on the bottom and the remaining variables in ascending order
2	Stacks the first group on the top and the remaining variables in descending order

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

CUTPOINT, MIDPOINT, and NINTERVAL: Session subcommands for specifying cutpoints and midpoints on page 1087

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot on page 1130

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

DOT: Session subcommand for displaying a symbol for each data value on page 1094

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

STEM-AND-LEAF: Session command for creating a stem-and-leaf plot

STEM-AND-LEAF C...C

Displays a separate stem-and-leaf plot for each C. Provides a display that is similar to a histogram, but uses the actual numbers to create the display.

BY C

Produces a separate display for each value in C. All displays for one column are put on the same scale. C must contain integers from –10,000 to +10,000 or the missing value code (*).

TRIM

Trims all values beyond the inner fences, and shows them on special lines labeled LO and HI. TRIM and BY cannot be used together.

INCREMENT K

Controls scaling. K is the distance between two lines of the display (that is, the distance between the smallest possible number on one line and the smallest possible number on the next).

PPLOT: Session command for creating a probability plot

PPLOT C...C

Displays a separate probability plot for each column. Minitab calculates the fitted cumulative distribution function (CDF) and associated confidence intervals based on parameters estimated from your data, or on historical parameters if you provide them using HPARAMETERS. Parameter estimates or historical parameters are displayed

in an output table along with an Anderson-Darling (AD) goodness-of-fit statistic and associated p-value, and the number of observations.

Note To change the method used to calculate the plot points, choose **File** > **Options** > **Individual Graphs** > **Probability Plots**.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

PTILES: Session subcommand for specifying the location of the percentile lines on a graph on page 1139

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

PERCENT

PERCENT specifies a percent scale for the y-axis.

PROBABILITY

PROBABILITY specifies a probability scale for the y-axis.

SCORES

SCORES specifies that the inverse cumulative probability from the standardized version of the chosen distribution be plotted on the y-axis.

UATHRESHOLD

UATHRESHOLD specifies that the x-axis will not be adjusted for the threshold parameter.

Distributions

NORMAL

Fits a normal distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

LNORMAL

Fits a lognormal distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

TLNORMAL

Fits a 3-parameter lognormal distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

GAMMA

Fits a gamma distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

TGAMMA

Fits a 3-parameter gamma distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

EXPONENTIAL

Fits a exponential distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

TEXPONENTIAL

Fits a 2-parameter exponential distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

SEXTREMEVALUE

Fits a smallest extreme value distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

WEIBULL

Fits a Weibull distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

TWEIBULL

Fits a 3-parameter Weibull distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

LEXTREMEVALUE

Fits a largest extreme value distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

LOGISTIC

Fits a logistic distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

LLOGISTIC

Fits a loglogistic distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

TLLOGISTIC

Fits a 3-parameter loglogistic distribution. If you don't use HPARAMETERS to specify historical parameters, Minitab estimates the mean and standard deviation from the data.

Data view

FITD: Session command for fitting a distribution to the data on a probability plot on page 1097

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

HPARAMETERS

Specifies the distribution parameters. If you want Minitab to estimate all parameters, do not use HPARAMETERS. Parameter arguments are given in the order shown in the table below. For each parameter that you want Minitab to estimate, enter ' * ' for K. For example:

- To have Minitab estimate the mean for a normal distribution, enter ' * ' for the mean and a value for the standard deviation.

```
NORMAL;
  HPARAMETERS   '*'   7.
```

- To have Minitab estimate the scale for a 3-parameter Weibull distribution, enter ' * ' for the scale and values for the shape and threshold.

```
TWEIBULL;
  HPARAMETERS   13   '*'   4.
```

Use the following table to determine the necessary parameters and restrictions.

Parameters	Distributions	Minitab can estimate
Mean and standard deviation	NORMAL	Mean and standard deviation Mean, given standard deviation
Location and scale	LOGNORMAL SEXTREME (smallest extreme value) LEXTREME (largest extreme value) LOGISTIC LOGLOGISTIC	Location and scale Location, given scale
Shape and scale	GAMMA Weibull	Shape and scale Scale, given shape
Mean	EXPONENTIAL	Mean
Scale and threshold	TEXPONENTIAL (2-parameter exponential)	Scale and threshold Scale, given threshold
Shape, scale, and threshold	TWEIBULL (3-parameter Weibull) TGAMMA (3-parameter gamma)	Shape, scale, and threshold Shape and scale, given threshold Scale and threshold, given shape Scale, given shape and threshold
Location, scale, and threshold	TLNORMAL (3-parameter lognormal) TLLOGISTIC (3-parameter loglogistic)	Location, scale, and threshold Location and scale, given threshold Location and threshold, given scale Location, given scale and threshold

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

ECDF: Session command for creating an empirical CDF plot

ECDF C...C

Displays a separate empirical CDF plot for each column.

An empirical cumulative distribution function (ECDF) resembles a cumulative histogram without bins. Minitab plots the value of each observation against the percentage of values in the sample that are less than or equal to that value. In this respect, an ECDF is similar to a probability plot except that both axes are linear, which can make interpreting the ECDF more intuitive.

Y-scale type

PERCENT

Specifies a percent scale for the y-axis.

PROBABILITY

Specifies a probability scale for the y-axis.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

PTILES: Session subcommand for specifying the location of the percentile lines on a graph on page 1139

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CONNECT: Session subcommand for connecting points with lines on page 1085

DISTRIBUTION: Session command for fitting a distribution on page 1091

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

DPLOT: Session command for creating a probability distribution plot

DPLOT

Creates a probability distribution plot.

Use to view and compare the shape of distribution curves and to view areas under distribution curves corresponding to either probabilities or data values. These graphs plot probability density functions (PDF) which describe the likelihood of each value that a variable can take on. Typically, you specify the distribution, parameter values, and, sometimes, an area under the curve.

DISTRIBUTION

Draws the distribution and parameters as specified in its subcommand. You can have an unlimited number of DISTRIBUTION subcommands per plot.

SHADE K

Adds a shaded region to the plot that corresponds to a probability or x-value. For more information and subcommands of SHADE, see the following section "Shaded area".

Shaded area

SHADE K

Shade a region of a single distribution corresponding to a probability or x-value using this subcommand of DISTRIBUTION. You can use this subcommand only once per DISTRIBUTION command. Specify the region to shade: right tail (K=1), left tail (K=2), equal tails (K=3), or middle (K=4).

SHTYPE K

Specifies shading by probability (K=1) or x-value (K=2).

SHVALUE K [K]

Specify one or two probabilities or x-values depending upon the values specified for SHADE and SHTYPE.

If SHADE = 1–3, then one value is required.

If SHADE = 4, then two values are required.

If SHTYPE = 1, then the value or values must be between 0 and 1 (noninclusive).

TYPE K

Specifies the fill type of the shading in K, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

Specifies the background color of the shading in K, which is a number from Numbers for colors to use in session commands on page 1255.

ETYPE K

Specifies the fill edge type of the shading in K, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K
Specifies the fill edge color of the shading in K, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K
Specifies the fill edge size, which can be any positive number.

Distributions

CHISQUARE K
Draws a Chi-square distribution with degrees of freedom = K.

NORMAL K K
Draws a normal distribution with mean = K and standard deviation = K.

F K K
Draws an F distribution with numerator degrees of freedom = K and denominator degrees of freedom = K.

T K
Draws a t distribution with degrees of freedom = K.

UNIFORM K K
Draws a uniform distribution with lower endpoint = K and upper endpoint = K.

BETA K K
Draws a beta distribution with first shape parameter = K and second shape parameter = K.

LNORMAL K K K
Draws a lognormal distribution with location = K, scale = K, and threshold (optional) = K.

BINOMIAL K K
Draws a binomial distribution with number of trials = K and event probability = K.

GAMMA K K K
Draws a gamma distribution with shape = K, scale = K, and threshold (optional) = K.

CAUCHY K K
Draws a Cauchy distribution with location = K and scale = K.

EXPONENTIAL K K
Draws an exponential distribution with scale = K, and threshold (optional) = K.

INTEGER K K
Draws an integer distribution with minimum value = K and maximum value = K.

SEXTREME K K
Draws a smallest extreme value distribution with location = K and scale = K.

WEIBULL K K K
Draws a Weibull distribution with shape = K, scale = K, and threshold (optional) = K.

LAPLACE K K
Draws a Laplace distribution with location = K and scale = K.

LEXTREME K K
Draws a largest extreme value distribution with location = K and scale = K.

LOGISTIC K K
Draws a logistic distribution with location = K and scale = K.

LLOGISTIC K K K
Draws a loglogistic distribution with location = K, scale = K, and threshold (optional) = K.

GEOMETRIC K
Draws a geometric distribution with event probability = K.

HYPERGEOMETRIC K K K
Draws a hypergeometric distribution with population size = K, event count = K, and sample size = K.

NEGBINOMIAL K K
Draws a negative binomial distribution with event probability = K and number of events needed = K.

POISSON K
Draws a Poisson distribution with mean = K.

TRIANGULAR K K K
Draws a triangular distribution with lower endpoint = K, mode = K, and upper endpoint = K.

DISCRETE C C
Draws a discrete distribution using values in C and probabilities in C.

NONEVENT
Models the number of non-events for the geometric and negative binomial distributions.

TOTAL
Models the number of trials for the geometric and negative binomial distributions.

Fill pattern and background fill color of the distribution bar for discrete distributions

BTYPE K
Specifies the fill pattern in K, which is a number from Numbers for fill types to use in session commands on page 1256.

BCOLOR K
Specifies the background color in K, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the fill edge line for discrete distributions

BETYPE K
Specifies the fill edge type in K, which is a number from Numbers for line types to use in session commands on page 1256.

BECOLOR K
Specifies the fill edge color, which is a number from Numbers for colors to use in session commands on page 1255.

BESIZE K
Specifies the fill edge size, which can be any positive number.

Line type, color, and size for continuous distributions

LTYPE K
Specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256.

LCOLOR K
Specifies the line color in K, which is a number from Numbers for colors to use in session commands on page 1255.

LSIZE K
Specifies the line size, which can be any number that is greater than 0.

Labels

POSITION K "K"
Specify a label to apply font, font color, font size, and label angle changes. The first K is the specific label; the second "K" is the optional label text. The positions are ordered left-to-right and top-to-bottom.

The label attributes and positioning commands below can be used as subcommands to POSITION. If POSITION is not used, or after executing ENDPOSITION, the attribute and positioning commands apply to all labels.

TFONT *"font name"*
TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

ENDPOSITION
End changes by label position.

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC

Specifies to display the label in italic font.

NOITALIC

Specifies to not display the label in italic font.

PLACEMENT K K

Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

SHOW K

Show (K=1) or do not show (K=0) the label specified by POSITION.

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Annotation

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130: NODFOOTNOTE does not work with DPLOT.

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TEXT: Session subcommand for displaying text on a graph on page 1161

TITLE: Session subcommand for adding a title to a graph on page 1162

Region

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region on page 1158

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

BOXPLOT: Session command for creating a boxplot

Use to examine characteristics of a sample distribution. Also called box-and-whisker plots.

The data must be numeric or date/time. The categorical grouping data can be numeric, date/time, or text.

BOXPLOT C...C

Displays a separate boxplot for each C.

The following command language displays two boxplots: one of Sales and one of Advertis.

```
BOXPLOT 'Sales' 'Advertis'.
```

BOXPLOT (C...C) * C

Displays a separate graph for each C on the left, with a box for each category of the C on the right.

The following command language displays a boxplot of yearly sales.

```
BOXPLOT 'Sales'*'Year'.
```

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot on page 1106

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot on page 1124

MEDLAB: Session subcommand for labeling medians on a boxplot on page 1128

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

OLAB: Session subcommand for labeling outliers on a boxplot on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CIBOX: The session subcommand for displaying a median confidence interval box on a boxplot on page 1078

CMEAN: Session subcommand for connecting means with lines on a boxplot on page 1083

CMEDIAN: Session subcommand for connecting means with lines on a boxplot on page 1084

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot on page 1109

IQRBOX: Session subcommand for displaying an interquartile range box on a boxplot on page 1114

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot on page 1126

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot on page 1127

OUTLIER: Session subcommand for displaying a symbol for each outlier on a boxplot on page 1133

PROPORTIONAL and NOPROPORTIONAL: Session subcommands for making the boxes on a boxplot proportional to the square root of the number of observations in the boxes on page 1139

RBOX: Session subcommand for displaying a range box on a boxplot on page 1143

SOFFSET: Session subcommand for offsetting points from the center on page 1155

WHISKER: Session subcommand for controlling the display of whiskers on a boxplot on page 1166

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

INTPLOT: Session command for creating an interval plot

INTPLOT C...C

INTPLOT (C...C) * C

Use to plot means and confidence intervals for one or more variables. An interval plot illustrates both a measure of central tendency and the variability of the data.

By default, Minitab displays confidence intervals, but you can change the display type to standard error bars using the INTBAR subcommand.

The data must be numeric or date/time. The categorical grouping data can be numeric, date/time, or text.

INTPLOT C...C displays a separate interval plot for each graph variable.

INTPLOT (C...C) * C displays a separate graph for each C on the left, with an interval for each category of the C on the right.

The following table shows a few of the possible ways to generate interval plots.

Command language	What the command language displays
`INTPLOT 'Sales' 'Advertis'.`	Two graphs: interval plots of Sales and of Advertis
`INTPLOT 'Sales'*'Year'.`	A graph with intervals representing yearly sales

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot on page 1106

INTLAB: Session subcommand for labeling interval bar endpoints on an interval plot on page 1112

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot on page 1124

MEDLAB: Session subcommand for labeling medians on a boxplot on page 1128

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

BAR: Session subcommand for representing data values with bars on page 1076

CMEAN: Session subcommand for connecting means with lines on a boxplot on page 1083

CMEDIAN: Session subcommand for connecting means with lines on a boxplot on page 1084

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot on page 1109

INTBAR: Session subcommand for displaying a vertical line with horizontal lines at the endpoints of the confidence interval for the mean on page 1110

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot on page 1126

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot on page 1127

SOFFSET: Session subcommand for offsetting points from the center on page 1155

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

INDPLOT: Session command for creating an individual value plot

INDPLOT C...C

INDPLOT (C...C) * C

Use to assess and compare sample distributions by plotting individual values for each variable or group in a vertical column, making it easy to spot outliers and see distribution shape. By default, Minitab adds horizontal jitter so that overlapping points are visible.

The data must be numeric or date/time. The categorical grouping data can be numeric, date/time, or text.

The table below shows a few of the possible ways to display interval plots.

INDPLOT C...C displays a separate individual value plot for each graph variable.

INDPLOT (C...C) * C displays a separate graph for each C on the left, with a column of individual values for each category of the C on the right.

The following table shows two of the possible ways to display individual value plots.

Command language	What the command language displays
`INDPLOT 'Sales' 'Advertis'.`	Two individual value plots: one of Sales and one of Advertis
`INDPLOT 'Sales'*'Year'.`	A individual value plot of yearly sales

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster on page 1101

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot on page 1106

INTLAB: Session subcommand for labeling interval bar endpoints on an interval plot on page 1112

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot on page 1124

MEDLAB: Session subcommand for labeling medians on a boxplot on page 1128

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

BAR: Session subcommand for representing data values with bars on page 1076

CMEAN: Session subcommand for connecting means with lines on a boxplot on page 1083

CMEDIAN: Session subcommand for connecting means with lines on a boxplot on page 1084

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot on page 1109

INTBAR: Session subcommand for displaying a vertical line with horizontal lines at the endpoints of the confidence interval for the mean on page 1110

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot on page 1126

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot on page 1127

SOFFSET: Session subcommand for offsetting points from the center on page 1155

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LPLOT: Session command for creating a line plot

Use to compare response patterns for two or more groups. The line plot has two distinct applications: looking for interactions in simple two-way designs and looking for consistency in response patterns of many groups. Each point on a line plot can represent a Bar chart functions on page 1216 of a category, or summary values from a table.

LPLOT *F*(C) * C

GROUP C

Displays a line plot of a function of a single response.

The following command language displays a line plot of mean hole diameter ('Diameter') as produced on two different production lines ('Line') each of which use two different drills ('Drill'). Drill type is identified through the legend.

```
LPLO MEAN('Diameter')*'Line';
 GROUP 'Drill'.
```

LPLOT *F*(C...C) * C

OVERLAY

Displays a line plot of a function of a multiple responses.

The following command language displays a line plot of the mean temperature ('Temp') and nutrient quotient of flour produced on two different types of grain mill ('Mill').

```
LPLO MEAN('Temp''Nutrient')*'Mill';
 OVERLAY.
```

LPLOT (C...C) * C

SUMMARIZED

Displays a line plot of tabled, multivariate data with one grouping column.

The following command language displays a line plot of the 25 chemical profiles (tabled in columns 'ChemProfile1'-'ChemProfile25') labeled by the component levels in 'Component.'

```
LPLO('ChemProfile1'-'ChemProfile25')*'Component';
 SUMMARIZED.
```

Plot options

PERCENT

Use a percent scale on the y-axis.

SLEVEL *1*

Use with a percent scale to display the percent within each x-category so that all points at the designated x-category total 100%. If you do not use SLEVEL 1, the line plot displays the percent across all categories.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CONNECT: Session subcommand for connecting points with lines on page 1085

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

GROUP: Session subcommand for specifying categorical variables for grouping on page 1102

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

PIECHART: Session command for creating a pie chart

Use to see the proportion of each data category relative to the whole data set.

PIECHART C...C

Displays a separate pie chart for each column of raw data (numeric, text, or date/time), where each row represents the category of a single observation. Minitab creates a slice for each unique value with the area representing the proportion of that value in the column.

PIECHART C*C...C*C

PIECHART C*C...C

Displays a pie chart of summarized data, where categories (or groups) are in the second set of columns and summarized data (numeric or date/time data that describe the category) are in the first set of columns. Minitab creates a pie chart for every combination of summary columns and categorical columns, using the worksheet order of the data.

Pie chart options

DONUT

Change display to a donut.

SEMICIRCLE

Change display to a semicircle.

COMBINE K

Combines categories with individual percentages less than or equal to K% into a category named Others (when one or more slices are at or below K%). When you use COMBINE, the Others category is always placed last, even when you specify INCREASING or DECREASING.

INCREASING

Orders slices from smallest to largest.

DECREASING

Orders slices from largest to smallest.

DIAMETER K

Specifies the pie chart diameter in figure units. K must be greater than 0 and less than or equal to 1.

START K

Specifies the starting angle in degrees for the slices. By default, slices start at 90° (12 o'clock) and go clockwise. When you specify START 0, slices start at the 3 o'clock position.

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SLABEL: Session subcommand for labeling the slices on a pie chart on page 1153

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

SLICE: Session subcommand for specifying the attributes of slices on a pie chart on page 1154

Multiple graphs

PANEL: Session subcommand for paneling graphs on page 1134

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

FREQUENCY: Session subcommand for using a frequency column for a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

TSPLOT: Session command for creating a time series plot

TSPLOT C...C

Displays a separate time series plot for each column.

Use to evaluate trends in data over time. Minitab plots time series data on the y-axis versus time on the x-axis. Minitab plots the data in worksheet order, in equally-spaced time intervals.

The data must be one or more columns of time series data (regular sequences of observations of a numeric or date/time variable over time).

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Time scale

For information about the following subcommands, go to Time unit subcommands on page 1263.

INDEX

STAMP

SECOND

MINUTE

HOUR

DAY

MONTH

QUARTER

YEAR

START

INCREMENT

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CONNECT: Session subcommand for connecting points with lines on page 1085

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot on page 1122

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

PARPLOT: Session command for creating a parallel plot

Use to compare observations or groups on multiple variables by plotting each observation on parallel coordinates.

PARPLOT C...C

Displays a parallel plot where each column (numeric), represents a separate y-scale on the plot. Every column must have the same number of rows.

Parallel plot options

RLABEL C

Specify row labels. A line is plotted for each row with a label in the legend. `RLABEL` C can be specified with `GROUP` C but not `MEAN` C.

GROUPC

Specify a column that groups the rows into categories. Minitab uses the group variable to create a legend for the parallel plot. `GROUP` C can be specified with `RLABEL` C but not `MEAN` C.

MEANC

Specify a column that groups the rows into categories. Minitab plots the mean for each group for each variable. Unique labels appear in legend. If `MEAN` C is specified then `RLABEL` C and `GROUP` C cannot be specified.

SORT

Sort the Y variables in descending order based on the variation between points plotted for each variable. If you specify `STANDARD`, the scale is sorted by the standard deviation of each variable. If you specify `SHARED`, the scale is sorted by the kurtosis of each variable. You cannot use this option when you specify `MINMAX`.

SCALE 2

Indicates the y scales. MINMAX, STANDARD, and SHARED are subcommands of SCALE 2.

MINMAX

Use to plot each variable with a unique y-scale. The minimum and maximum values on each scale are parallel for all of the columns. A row that contains all of the minimum or maximum values for each variable will be a horizontal line. You cannot use MINMAX with STANDARD or SHARED.

STANDARD

Use to plot each variable with a unique y-scale. The minimum and maximum values for each column are converted to the transformed standard deviation units, and the scales are based on these values. You cannot use STANDARD with MINMAX or SHARED.

SHARED

Use to plot a single y-scale for every variable. The minimum and maximum values for the scale are the overall minimum and maximum values from all the data that you entered. You cannot use SHARED with STANDARD or MINMAX.

MULTIPLE C...C

Creates separate graphs for each group created by the grouping columns C...C. The grouping columns can be numeric or text, but must be the same length as the columns that you specify with PARPLOT. Minitab treats missing values as a separate group.

INDEPENDENT

Creates separate plots for each group created by the individual columns specified with MULTIPLE independent of the other columns. For example, suppose you specify two columns with MULTIPLE. The first column contains the gender, either male or female, and the other contains the marriage status, either married or single. If you specify INDEPENDENT, Minitab creates a graph for each group created by the columns individually:

- Male
- Female
- Single
- Married

If you do not specify INDEPENDENT, Minitab creates a graph for each group created by all the columns:

- Male and Single
- Male and Married
- Female and Single
- Female and Married

ARGRAPH: Session command for creating an area graph

ARGRAPH C...C

Use to evaluate trends in data over time by plotting series of sequential observations. An area graph stacks areas for each time series, illustrating the proportion each series contributes to the sum.

You need one or more columns of time series data (regular sequences of observations of a numeric or date/time variable over time) of equal length.

ORDER

Stacks areas by the degree of variation, with the area having the smallest standard deviation on the bottom.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TSHOW: Session subcommand for specifying the level of tick labels that are displayed on page 1164

Time scale

Time unit subcommands on page 1263

Labels

DATLAB: Session command for labeling data values on page 1089

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

Multiple graphs

PANEL: Session subcommand for paneling graphs on page 1134

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

GAPS: Session subcommand for displaying a gap in time on a graph on page 1100

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

CONTOURPLOT: Session command for creating a contour plot

In a contour plot, values for the x- and y-variables are represented on the x- and y-axes, respectively. Values for a third variable, z, are estimated and represented by contour lines and/or shaded regions. A contour plot is like a topographical map in which x-, y-, and z-values are plotted instead of longitude, latitude, and altitude.

Note You can create mesh data for CONTOURPLOT with MESH on page 85.

CONTOURPLOT C * C * C

Displays a contour plot of z (1st column) as a function of y (2nd column) and x (3rd column).

CONTOURPLOT (C...C) * C * C

Displays contour plots for each z as a function of the same y and x.

Contour options

APOLYNOMIAL

Specifies the Akima's polynomial method to estimate the z-values for irregular data (data not on a rectangular grid). If x-y data are sampled on a rectangular grid, APOLYNOMIAL is ignored. APOLYNOMIAL and DISTANCE are mutually exclusive. The last one entered will be used.

BOUNDARY K

BOUNDARY sets the z-estimate for points far from your x-y data.

DISTANCE

Specifies the Distance method to estimate the z-values for irregular data (data not on a rectangular grid). If data are sampled on a rectangular grid, DISTANCE is ignored. APOLYNOMIAL and DISTANCE are mutually exclusive. The last one entered will be used.

NOSTANDARDIZE

Specifies not to standardize x- and y-data before interpolation. Use if your data are not measured on different scales.

POWER K

POWER specifies the amount of local variation smoothing. K is an integer greater than 0 and less than or equal to 12. The closer to 0, the closer the fits are to the overall mean. The farther from 0, the more weight is placed on local variation.

LEVELS K...K

LEVELS C

Specifies from 2 to 11 z-values for the contour lines. If you specify LEVELS more than once, CONTOURPLOT uses only the last occurrence.

NLEVELS K

Specifies the number of contour lines as an integer from 2 to 11. If you specify NLEVELS more than once, CONTOURPLOT uses only the last occurrence.

NMESH

Specifies the number of mesh values in the x- or y-direction. For example, `NMESH 1 20` specifies 20 mesh positions in the x-direction and `NMESH 2 30` specifies 30 positions in the y-direction.

By default, Minitab uses the mesh defined by your data if they are regular, or a 15 by 15 mesh for irregular data. (Data with a regular shape form a grid with evenly spaced intervals. Data with an irregular shape are not located on evenly spaced intervals.)

If you specify NMESH more than once for x or y, CONTOURPLOT uses only the last occurrence.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

CONNECT: Session subcommand for connecting points with lines on page 1085

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Data options

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

PLTX: Session command for creating a 3D scatterplot

Use to evaluate relationships between three variables at once by plotting data on three axes: z, y, and x.

PLTX with no subcommands displays symbols. Use SYMBOL if you want symbols to be displayed with PROJECT.

PLTX C * C * C

Displays a three-dimensional scatterplot of z (1st column) by y (2nd column) by x (3rd column).

PLTX (C...C) * C * C

Displays a separate three-dimensional scatterplot for each z by the same y and x.

Regions

DATA [K K K K]

DATA for PLTX has limited functionality compared to DATA for other commands. DATA specifies the position of the box in terms of the x-minimum, x-maximum, y-minimum, and y-maximum as values from 0 to 1. Subcommands of DATA are ignored for PLTX.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points on page 1116

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

3D orientation

VASPECT: Session subcommand for specifying the aspect ratio of the data box on page 1164

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph on page 1165

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units on page 1165

VPOSITION: Session subcommand for specifying the view position as a ratio on page 1165

VUP: Session subcommand for specifying which direction is up in relation to the data box on page 1165

Data options

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

DATA: Session subcommand for controlling the data region within the figure region on page 1088

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

SURFACEPLOT: Session command for creating a surface plot

Use to evaluate relationships between three variables at one time by plotting data on three axes: z, y, and x. Like a three-dimensional scatterplot, a three-dimensional surface plot has three axes. In addition, SURFACEPLOT uses interpolation to produce a continuous surface (surface plot) or grid (wireframe plot) of z-values that fits your data. For information on a three-dimensional scatterplot, go to PLTX on page 1068.

SURFACEPLOT with no subcommands displays a fitted surface. Use SFIT if you want a fitted surface or wireframe to be displayed with PROJECT or SYMBOL.

Note You can create mesh data for SURFACEPLOT with MESH on page 85.

SURFACEPLOT C * C * C

Displays a three-dimensional surface plot of z (1st column) by y (2nd column) by x (3rd column).

SURFACEPLOT (C...C) * C * C
Displays a separate surface plot for each z by the same y and x.

SURFACEPLOT C * C * C

WIREFRAME
Displays a three-dimensional wireframe plot of z (1st column) by y (2nd column) by x (3rd column).

SURFACEPLOT (C...C) * C * C

WIREFRAME
Displays a separate wireframe plot for each z by the same y and x.

Surface options

APOLYNOMIAL
Specifies the Akima's polynomial method to estimate the z-values for irregular data (data not on a rectangular grid). If x-y data are sampled on a rectangular grid, APOLYNOMIAL is ignored. APOLYNOMIAL and DISTANCE are mutually exclusive. The last one entered will be used.

BOUNDARY K
BOUNDARY sets the z-estimate for points far from your x-y data.

DISTANCE
Specifies the Distance method to estimate the z-values for irregular data (data not on a rectangular grid). If data are sampled on a rectangular grid, DISTANCE is ignored. APOLYNOMIAL and DISTANCE are mutually exclusive. The last one entered will be used.

NOSTANDARDIZE
Specifies not to standardize x- and y-data before interpolation. Use if your data are not measured on different scales.

POWER K
POWER specifies the amount of local variation smoothing. K is an integer greater than 0 and less than or equal to 12. The closer to 0, the closer the fits are to the overall mean. The farther from 0, the more weight is placed on local variation.

LBRIGHT K
Specifies the brightness (as an integer from 0 to 100) of the lights illuminating the surface. LBRIGHT has no effect when WIREFRAME is used.

LIGHT K K K K
The first argument for LIGHT specifies the number of the light (1, 2, or 3). The remaining arguments specify the position of the light in x-, y-, and z-coordinates. For example LIGHT 1 0 0 1 specifies that light 1 is to be centered on the x- and y-axes, one unit above the surface in the positive z-direction. LIGHT has no effect when WIREFRAME is used.

COLOR K
COLOR specifies the color of the light as an integer. To find the integer for a color, go to Numbers for colors to use in session commands on page 1255.

VISIBILITY K
VISIBILITY specifies whether the light is on (0) or off (1).

NMESH

Specifies the number of mesh values in the x- or y-direction. For example, `NMESH 1 20` specifies 20 mesh positions in the x-direction and `NMESH 2 30` specifies 30 positions in the y-direction.

By default, Minitab uses the mesh defined by your data if they are regular, or a 15 by 15 mesh for irregular data. (Data with a regular shape form a grid with evenly spaced intervals. Data with an irregular shape are not located on evenly spaced intervals.)

If you specify NMESH more than once for x or y, SURFACEPLOT uses only the last occurrence.

WIREFRAME

Specifies a wireframe display.

Scale

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

PROJECT: Session subcommand for extending projection lines from each point to the x-axis on page 1137

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot on page 1152

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

3D orientation

VASPECT: Session subcommand for specifying the aspect ratio of the data box on page 1164

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph on page 1165

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units on page 1165

VPOSITION: Session subcommand for specifying the view position as a ratio on page 1165

VUP: Session subcommand for specifying which direction is up in relation to the data box on page 1165

Data options

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph on page 1109

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graph Options

AREA: The session subcommand for shading the area below the data values to the base

AREA [C...C]

Shades the area below the data values to the base (scatterplot, matrix plot, histogram, or area graph) or the areas between contours (contour plot), where C...C can specify up to 3 grouping variables.

For grouping variables specified on the main command line:

- You can use up to three grouping columns.
- Grouping variables are not allowed with CONTOURPLOT and ARGRAPH.
- The AREA arguments are grouping variables. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different area attributes.

```
PLOT C1*C2;
  AREA C2.
```

For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266 for more information.

DIRECTION K

Specifies the direction the area extends away from the data values. DIRECTION is not available with HISTOGRAM.

Value of K	Direction
1	Horizontally to the base
2 (default)	Vertically to the base

Area edge line

Specifies the type, color, and width of the area edge line.

ETYPE K...K

ETYPE C

ETYPE specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each area or store the line types in C.

ETYPE does not apply for CONTOURPLOT.

ECOLOR K...K

ECOLOR C

ECOLOR specifies the line color in K, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each area or store the line colors in C.

ECOLOR does not apply for CONTOURPLOT.

ESIZE K...K

ESIZE C

ESIZE specifies the line width, which can be any positive number. Enter one K for each area or store the line sizes in C.

ESIZE does not apply for CONTOURPLOT.

Area

Specifies the fill pattern and background fill color of the area.

TYPE K...K

TYPE C

TYPE specifies the fill pattern in K, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each area or store the fill types in C.

COLOR K...K

COLOR C

COLOR specifies the background color in K, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each area or store the fill colors in C.

Color ramp for the levels of a contour plot

Specifies the color ramp to be used for the levels of a contour plot.

RAMP K [K]

RAMP specifies a one-color ramp or a two-color ramp. Use one argument for a one-color ramp and two arguments for a two-color ramp for example. The constant arguments, K, refer to the ramp (1=red, 2=green, 3=blue, 4=gray). The first argument is used for the high values, and the second is used for the low values. One-color ramps, available only with contour plots with five contour levels or less, use light shades to represent the low values and dark shades to represent the high values.

SPECTRUM

SPECTRUM specifies the spectrum color ramp, which mimics the spectrum with red representing the low values and blue representing the high values.

Data point connect lines

STRAIGHT (default)

STRAIGHT connects data points with straight lines.

STEP [K]

The following table shows how STEP connects data points. STEPS is not available with CONTOURPLOT, ARGRAPH, or HISTOGRAM.

Value of K	How STEP connects data points
–1	Makes steps with each data point at the left of a step level.
0 (default)	Puts each point in the middle of the step level.
1	Makes steps with each point at the right of the step level.

AXLABEL: Session subcommand for customizing graph axis labels

AXLABEL K *"string"*

AXLABEL K C

Generates customized axis labels. You can also modify axis lines and labels using graph editing, which may be easier than using AXLABEL.

Specifies the axis to be affected and its label.

Value of K	Axis
1	X-axis
2	Y-axis
3	Z-axis on a 3D graph

You can enter a custom axis label in quotation marks or specify a column that contains only the axis label that you want to use.

MODEL K

Specifies the graph that will accept the axis label subcommand. Use this subcommand only when you create multiple graphs with a single command. Allows you to apply its related subcommand and subsubcommands to a specific graph. The argument refers to the order in which the graphs are created. `MODEL 1` affects the first graph created, `MODEL 2` the second, and so on.

SECSCALE

Specifies that the scale is a secondary scale. Invoking this subsubcommand associates its related subcommand with the secondary scale.

ADISPLAY K [K...K]

Specifies the side and display of the axis label.

Value of K	How the axis label is displayed
0	Hides the axis label
1	Displays the axis label on the low side
2	Displays the axis label on the high side

LSHOW K

Shows or hides axis labels on a multi-level axis. By default, axis labels for all levels are shown. When you use LSHOW, only specified levels are shown. A K value of 1 refers to the first or outermost variable level, 2 to the second, and so on.

LABEL [K] *"string"*

Use LABEL in two settings: with matrix plots and with multi-level scales, such as those on nested boxplots.

With matrix plots, use LABEL K "string". The first argument specifies the variable number that will take the custom label that follows it in quotes. The first variable listed in the MATRIXPLOT command is referenced by 1, the second by 2, and so on.

With multi-level axes, use LABEL "string" with the subsubcommand ALEVEL K.

ALEVEL K

With multi-level axes, use LABEL "string" with the subsubcommand ALEVEL K. The argument to ALEVEL specifies the variable level that takes the custom label that follows the LABEL command in quotes. A K value of 1 refers to the first or outermost label, 2 to the second, and so on.

TFONT *"font"*

TFONT specifies the font and uses the font name as an argument.

TCOLOR K

TCOLOR specifies the color of the text in K, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

Use TSIZE or PSIZE to specify the size of the text. TSIZE takes values in Minitab units.

PSIZE K

Use TSIZE or PSIZE to specify the size of the text. PSIZE takes point units.

ANGLE K

ANGLE rotates the text K degrees in a counterclockwise direction around the reference point on the axis.

OFFSET K K

Moves the text from the default rotation of 0 0. `OFFSET 0 0` gives you axis labels directly on the axis line, centered at the midpoint. The first argument is the x-direction offset (−1 is to the left of the point and 1 is to the right). The second argument is the y-direction offset (−1 is below the point and 1 is above). Both are in figure units.

BOLD
Specifies that the axis labels are in bold.

ITALIC
Specifies that the axis labels are in italics.

UNDERLINE
Specifies that the axis labels are underlined.

NOBOLD
Specifies that the axis labels are not in bold.

NOITALIC
Specifies that the axis labels are not in italics.

NOUNDERLINE
Specifies that the axis labels are not underlined.

BAR: Session subcommand for representing data values with bars

BAR [C...C]
Displays a bar for each data value. If you use other data displays with a graph that has BAR as a default, such as histogram, you must issue the BAR subcommand.

For grouping variables specified on the main command line:

- Histograms can use up to three grouping columns.
- Bar chart, interval plot, and individual value plot can use up to four grouping columns.
  - For interval plot and individual value plot, grouping columns must be categorical variables that are included on the main graph command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different bar attributes.

    ```
    INTPLOT C1*C2;
    BAR C2.
    ```
  - For bar chart, grouping columns must be categorical variables that are GROUP, STACK, or, CLUSTER variables. In the following example, C3 is the grouping column, and, for each unique value in C3, Minitab assigns different bar attributes.

    ```
    CHART C1;
      GROUP C2 C3;
     BAR C3.
    ```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN K...K

VASSIGN C
For overlaid bar charts, interval plots, and individual value plots, assigns different attributes (color, fill type) to the bars representing each graph variable and displays a legend. (CHART, INTPLOT, and INDPLOT only)

BASE K...K

BASE C

Specifies the base position for each bar. (For more information, go to Base position for project lines, area, and bar on page 1217.) Bars extend from the data value to the base. The default base is the minimum data value for the y-scale.

Enter one K for each base or store the bases in C. If one value is specified it applies to all points.

You can draw a base line by using REFERENCE on page 1146.

Bar edge line

Specifies the type, color, and width of the bar edge line.

ETYPE K...K

ETYPE C

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each bar or store the line types in C.

ECOLOR K...K

ECOLOR C

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

ESIZE K...K

ESIZE C

ESIZE specifies the line width, which can be any positive number. Enter one K for each bar or store the line sizes in C.

Fill pattern and background fill color

Specifies the fill pattern and background fill color of the bar.

TYPE K...K

TYPE C

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each bar or store the fill types in C.

COLOR K...K

COLOR C

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the fill colors in C.

Positions where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, see Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

CIBOX: The session subcommand for displaying a median confidence interval box on a boxplot

CIBOX [C...C]

Displays a median confidence interval box on a boxplot. The box endpoints correspond to the values of a confidence interval for the median based on the sign test.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different box attributes.

```
BOXPLOT C1*C2;
    MEAN C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN K...K

VASSIGN C

For overlaid boxplots, assigns different attributes (color, fill type) to the median confidence interval boxes representing each graph variable and displays a legend.

Fill pattern and background fill color of the box

Specifies the fill pattern and background fill color of the box.

TYPE K...K

TYPE C

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each bar or store the fill types in C.

COLOR K...K

COLOR C

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the fill colors in C.

Type, color, and width of box edge line

ETYPE K...K

ETYPE C

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each bar or store the line types in C.

ECOLOR K...K

ECOLOR C

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

ESIZE K...K

ESIZE C

ESIZE specifies the line width, which can be any positive number. Enter one K for each bar or store the line sizes in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

CLIMITS: Session command for specifying attributes for control limit lines

CLIMITS

Specifies attributes for the upper and lower control limit lines and their labels on a control chart.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number that is specified in Numbers for line types to use in session commands on page 1256 list. Enter one K for line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number that is specified in Numbers for colors to use in session commands on page 1255 list. Enter one K for each line or store the line colors in C.

SIZE K...K

SIZE C

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Label attributes

LSHOW K

Specifies whether to display the label.

Value of K	Label
0	Do not display the label.
1 (default)	Display the label.

POSITION K

Specifies a particular control limit label. For example, `POSITION 3` refers to the upper control limit of the 3rd stage. The lower control limit position depends on the number of stages. For example, if there are 4 stages, then 1–4 refers to upper control limits and 5–8 refers to lower control limits. Use any of the subcommands, TFONT through ANGLE, to modify the label, then issue ENDPOSITION.

ENDPOSITION

Specifies the end of the POSITION subcommand language.

TFONT *"font name"*

TFONT specifies the text font. Enclose the font name "K" in double quotation marks.

TCOLOR K

TCOLOR specifies the label color and uses integer arguments that are specified in Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

BOLD

Specifies to display the label in bold font.

NOBOLD

Specifies to not display the label in bold font.

UNDERLINE

Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Label positioning

OFFSET K K
Moves the label. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies how far to move the title vertically. The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 specifies horizontal text.

CLINE: Session command for specifying attributes for a center line

CLINE
Specify attributes for the center line on a control chart. The CLINE attributes apply to all center lines on a chart.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C
TYPE specifies the line type, which is a number that is specified in Numbers for line types to use in session commands on page 1256 list. Enter one K for line or store the line types in C.

COLOR K...K

COLOR C
COLOR specifies the line color, which is a number that is specified in Numbers for colors to use in session commands on page 1255 list. Enter one K for each line or store the line colors in C.

SIZE K...K

SIZE C
SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Label attributes

LSHOW K
Specifies whether to display the label.

Value of K	Label
0	Do not display the label.
1 (default)	Display the label.

TFONT *"font name"*
TFONT specifies the text font. Enclose the font name "K" in double quotation marks.

TCOLOR K
TCOLOR specifies the label color and uses integer arguments that are specified in Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Label positioning

OFFSET K K
Moves the label. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies how far to move the title vertically. The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 specifies horizontal text.

CMEAN: Session subcommand for connecting means with lines on a boxplot

CMEAN

Connects means with lines on a boxplot, interval plot, or individual value plot. Mean lines are visible only if you have categorical variables in the boxplot.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main graph command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different line attributes.

```
BOXPLOT C1*C2;
  CMEAN C2.
```

- When used without grouping variables, connects the means of the boxes from left to right.

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN K...K

VASSIGN C

For overlaid graphs, connects mean symbols representing the same graph variable and displays a legend.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

SIZE K

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Lines that connect means

STRAIGHT (default)

Connects means with straight lines.

STEP K

Connects points with stepwise lines of the following possible justifications.

Value of K	Steps and justification
–1	Steps with each point at the left of a step level
0 (default)	Steps centered between the points
1	Steps with each point at the right of the step level

CMEDIAN: Session subcommand for connecting means with lines on a boxplot

CMEDIAN [C...C]

Connects medians with lines on a boxplot, interval plot, or individual value plot. Median lines are only visible if you have categorical variables in the boxplot.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main graph command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different line attributes.

```
BOXPLOT C1*C2;
  CMEDIAN C2.
```

- When used without grouping variables, connects the means of the boxes from left to right.

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN K...K

VASSIGN C

For overlaid graphs, connects median symbols representing the same graph variable and displays a legend.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

SIZE K

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Lines that connect means

STRAIGHT (default)
Connects medians with straight lines.

STEP K
Connects points with stepwise lines of the following possible justifications.

Value of K	Steps and justification
−1	Steps with each point at the left of a step level
0 (default)	Steps centered between the points
1	Steps with each point at the right of the step level

CONNECT: Session subcommand for connecting points with lines

CONNECT [C...C]
Connects points with lines, and turns off default bars, boxes, or symbols. CONNECT also restores connect lines when another data display command turns them off (so you can display for example, CONNECT and SYMBOL).

For grouping variables specified on the main command line:

- Can use up to three grouping columns.
- Bar chart can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main CHART command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different line attributes.

```
CHART C1;
  CONNECT C2.
```

- Grouping variables are not allowed with CONTOURPLOT and CONNECT.

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN
For overlaid bar charts, connects symbols representing the same graph variable and displays a legend.

ORDER K
Specifies the order in which points are connected. You have the following three choices:

Value of K	Order in which points are connected
0	Worksheet order (the order of occurrence in the worksheet)
1 (default)	X-order (from smallest to largest horizontally)
2	Y-order (from smallest to largest vertically)

You cannot use `ORDER 0` and STEP together.

Used only with PLOT and MATRIXPLOT.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

SIZE K

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Labels

NOLABEL

Does not show labels.

LABEL

Shows labels.

TYPE K...K

TYPE C

TYPE specifies the fill type, which is a number from Numbers for line types to use in session commands on page 1256.

COLOR K...K

COLOR C

COLOR specifies the background color for fill, which is a number from Numbers for colors to use in session commands on page 1255.

ETYPE K

ETYPE specifies the edge type for fill, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the edge color for fill, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the edge size for fill. The edge size can be any positive real number value.

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Lines that connect data values

STRAIGHT (default)
Connects values with straight lines.

STEP K
Connects points with stepwise lines of the following possible justifications.

Value of K	Steps and justification
−1	Steps with each point at the left of a step level
0 (default)	Steps centered between the points
1	Steps with each point at the right of the step level

CUTPOINT, MIDPOINT, and NINTERVAL: Session subcommands for specifying cutpoints and midpoints

You can specify your own midpoints or cutpoints by putting them on the subcommand line. Specify arguments in order, from smallest to largest. MIDPOINT requires equally spaced intervals. CUTPOINT does not.

CUTPOINT K...K

CUTPOINT C

Used without arguments, CUTPOINT places ticks at default cutpoint. To specify custom cutpoints, list each value (smallest to largest), create a patterned range, or specify a column containing the values. For example, the patterned range 10:50/5 places cutpoints at every 5 units from positions 10 to 50.

MIDPOINT K...K

MIDPOINT C

Used without arguments, MIDPOINT places ticks at default midpoint intervals. To specify custom midpoints, list each value (smallest to largest), create a patterned range, or specify a column containing the values. For example, the patterned range 10:50/5 places midpoints at every 5 units from positions 10 to 50.

NINTERVAL K [K]

NINTERVAL is a subsubcommand for CUTPOINT and MIDPOINT.

NINTERVAL with a single argument (between 2 and 1000, inclusive) allows you to specify a number of binning intervals to display on the x-axis. Minitab attempts to give useful tick values to histograms or dotplots produced with NINTERVAL. If you specify NINTERVAL with two arguments—for example, `NINTERVAL 3 7`—the graph has useful tick locations and between 3 and 7 intervals. Midpoints are displayed by default when you specify NINTERVAL. You can specify CUTPOINT or MIDPOINT without arguments and an NINTERVAL subcommand to display a graph with a certain number (or range) of intervals and cutpoints or midpoints. With NINTERVAL, the graph does not display first and/or last intervals that do not contain any observations.

DATA: Session subcommand for controlling the data region within the figure region

DATA [K K K K]

Controls size and placement of the data region within the figure region, the fill type and color of the data region, and the type, color and thickness of the edge line around the data region.

Represents the x-minimum, x-maximum, y-minimum, and y-maximum values for the data region, respectively, and uses figure units that range from 0 to 1. When used without arguments, the data region maintains its default size.

Fill type and color of the data region

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the data region border

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number.

DATLAB: Session command for labeling data values

DATLAB [C]

Labels data values. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  If graph data match worksheet rows, as for scatterplot, Minitab matches the label column to the data column by worksheet row number. If graph data are summaries of worksheet values, such as for histogram and bar chart, the label column order is left-to-right on the graph. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the ROWNUM subcommand to label data values with their worksheet row number.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands affect only the individual label until ENDPOSITION is issued.

YVALUE

Labels each data value with its y-axis value.

ROWNUM

Labels each data value with its worksheet row number. You cannot label histograms or bar charts with row numbers.

Labeling individual data values

POSITION K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION
After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*
TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	–1	0	1
1			

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

DISTRIBUTION: Session command for fitting a distribution

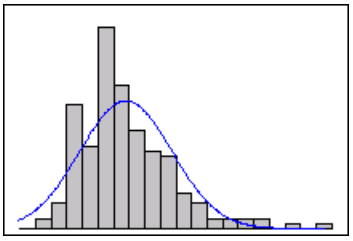
DISTRIBUTION [C...C]

Fits a distribution to the data, where C...C can specify up to 3 grouping variables. When you use grouping variables, DISTRIBUTION displays a separate distribution for each group.

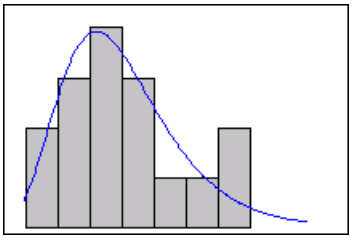
Fits a distribution to a histogram and empirical cdf graph to help you determine whether the data can be adequately modeled by the selected distribution. Minitab estimates the distribution parameters from the data unless you use the HPARAMETERS subcommand. Use the HPARAMETERS subsubcommand to enter 1–3 historical parameters.

Minitab provides 14 distributions: normal, lognormal, three-parameter lognormal, gamma, three-parameter gamma, exponential, two-parameter exponential, smallest extreme value, Weibull, three-parameter Weibull, largest extreme value, logistic, loglogistic, and three-parameter loglogistic.

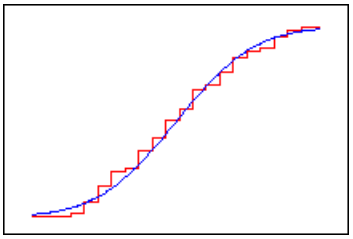
Histogram with a normal distribution



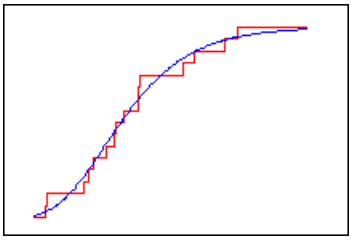
Histogram with a lognormal distribution



ECDF graph with a normal distribution



ECDF graph with a lognormal distribution



Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

SIZE K

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Distributions and parameters

NORMAL

Fits a normal distribution.

LNORMAL

Fits a lognormal distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

TLNORMAL

Fits a three-parameter lognormal distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

GAMMA

Fits a gamma distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

TGAMMA
Fits a three-parameter gamma distribution. Minitab estimates the shape, scale, and threshold from the data.

EXPONENTIAL
Fits a exponential distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

TEXPONENTIAL
Fits a two-parameter exponential distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

SEXTREMEVALUE
Fits a smallest extreme value distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

WEIBULL
Fits a Weibull distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

LEXTREMEVALUE
Fits a largest extreme value distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

LOGISTIC
Fits a logistic distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

LLOGISTIC
Fits a loglogistic distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

TLLOGISITIC
Fits a three-parameter loglogistic distribution. If you do not use HPARAMETERS to specify historical parameters, Minitab estimates them from the data.

HPARAMETERS K

HPARAMETERS K K

HPARAMETERS K K K
Specifies the distribution parameters. Parameter arguments are given in the order shown in the table below. For each parameter that you want Minitab to estimate, enter ' *' for K. For example:

- To have Minitab estimate the mean for a normal distribution, enter ' * ' for the mean and a value for the standard deviation.

```
DISTRIBUTION;
  NORMAL;
HPARAMETERS '*' 7.
```

- To have Minitab estimate the scale for a 3-parameter Weibull distribution, enter ' * ' for the scale and values for the shape and threshold.

```
DISTRIBUTION;
  TWEIBULL;
HPARAMETERS 13 '*' 4.
```

If you want Minitab to estimate all parameters, do not use HPARAMETERS.

Use the following table to determine the necessary parameters and restrictions.

Distribution subsubcommand	Parameters	What Minitab can estimate
NORMAL	Mean	Mean and standard deviation
	Standard deviation	Mean, given standard deviation
LOGNORMAL	Location	Location and scale
SEXTREME	Scale	Location, given scale
LEXTREME		
LOGISTIC		
LOGLOGISTIC		
GAMMA	Shape	Shape and scale
WEIBULL	Scale	Scale, given shape
EXPONENTIAL	Mean	Mean
TEXPONENTIAL	Scale	Scale and threshold
	Threshold	Scale, given threshold
TWEIBULL	Shape	Shape, scale, and threshold
TGAMMA	Scale	Shape and scale, given threshold
	Threshold	Scale and threshold, given shape
		Scale, given shape and threshold
TLNORMAL	Location	Location, scale, and threshold
TLLOGISTIC	Scale	Location and scale, given threshold
	Threshold	Location and threshold, given scale
		Location, given scale and threshold

DOT: Session subcommand for displaying a symbol for each data value

DOT [C...C]

Displays a symbol for each data value, or for multiple data values if there is insufficient room for individual symbols, on a dotplot, where C...C can specify up to 4 grouping variables.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.

- Grouping columns must be categorical variables that are included on the main DOTPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

```
DOTPLOT C1*C2;
  DOT C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid dotplots, assigns different attributes (color, type) to the dots representing each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C

TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph

ELLIPSE K K K K [K K K K]

ELLIPSE displays an ellipse at the points specified in the subcommand.

You can also put an ellipse anywhere on a graph using graph editing. Using the graph editing features is often easier, unless you want to place an ellipse in exact locations using predefined points.

An ellipse is defined by either four or eight arguments. If the ellipse is a circle, use four arguments (X Y X Y) that represent coordinates for the lower left and upper right corners of the square that defines the circle. Use eight arguments to represent coordinate of all four corners (lower left, lower right, upper right, upper left) of the rectangular region that defines a non-circular ellipse.

UNIT K

To place ellipses using figure units, use the argument 0. To specify data units, use 1.

With paneled graphs (including matrix plots), you can still place ellipses using data units if you also specify the panel with the UNIT command. UNIT 1 for example places the ellipse relative to the first panel of the graph and reads the ELLIPSE arguments in the data units of the first panel. You can associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, ELLIPSE always uses graph units.

Fill pattern and background fill color

Specifies the fill pattern and background fill color of the box.

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the line

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number. Enter a value that is greater than 0.

FIGURE: Session subcommand for controlling the figure region within the graph region

FIGURE [K K K K]

Controls size and placement of the figure region within the graph region, the fill type and color of the figure region, and the type, color and thickness of the edge line around the figure region.

The arguments represent the x-minimum, x-maximum, y-minimum, and y-maximum values for the figure region, respectively, and use graph units which range from 0 to 1. When used without arguments, the figure region maintains its default size.

Fill type and color of the figure region

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background fill color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the figure region border

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K.

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number.

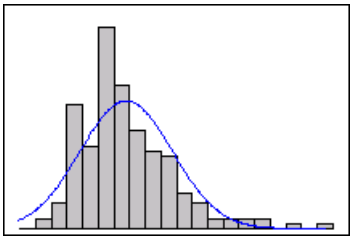
FITD: Session command for fitting a distribution to the data on a probability plot

FITD [C...C]

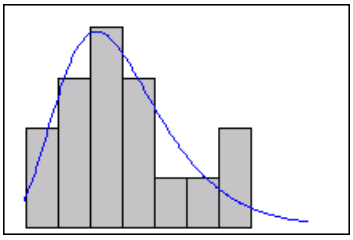
Fits a distribution to the data on a probability plot to help you determine whether the data can be adequately modeled by the selected distribution. To specify the distribution, go to PPLOT on page 1036.

Minitab provides 14 distributions: normal, lognormal, three-parameter lognormal, gamma, three-parameter gamma, exponential, two-parameter exponential, smallest extreme value, Weibull, three-parameter Weibull, largest extreme value, logistic, loglogistic, and three-parameter loglogistic.

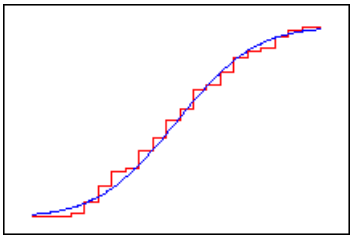
Histogram with a normal distribution



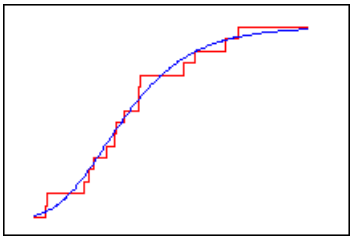
Histogram with a lognormal distribution



ECDF graph with a normal distribution



ECDF graph with a lognormal distribution



CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

NOCI

Hides the confidence interval.

Line type, color, and size

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each distribution fit or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each distribution fit or store the line colors in C.

SIZE K...K

SIZE C

SIZE specifies the line width, which can be any positive number. Enter one K for each distribution fit or store the line sizes in C.

FOOTNOTE: Session subcommand for adding a footnote to a graph

FOOTNOTE *["footnote"]*

Adds a footnote to a graph, placed in the lower left corner of the figure region by default.

- If you use FOOTNOTE more than one time, each additional footnote goes below the preceding footnote.

- For LAYOUT, the footnote goes at the lower left hand corner of the page, instead of the lower left hand corner of the figure region.
- You can suppress the default footnote with the subcommand NODFOOTNOTE.

Attributes

TFONT *"font name"*
TFONT specifies the font name for the footnote.

TCOLOR K
TCOLOR specifies the footnote text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the footnote in bold font.

NOBOLD
Specifies to not display the footnote in bold font.

UNDERLINE
Specifies to underline the footnote.

NOUNDERLINE
Specifies to not underline the footnote.

ITALIC
Specifies to display the footnote in italic font.

NOITALIC
Specifies to not display the footnote in italic font.

Positioning

LEFT (default)
Puts the footnote in the left of the figure region.

CENTER
Puts the footnote in the center of the figure region.

RIGHT
Puts the footnote in the right of the figure region.

OFFSET K K
Moves a footnote away from the position specified with LEFT, CENTER or RIGHT. The first argument specifies how far to move the footnote horizontally (in figure units). The second argument specifies the how far to move the footnote vertically (in figure units). The default offset is 0.0 0.0.

FREQUENCY: Session subcommand for using a frequency column for a graph

FREQUENCY C...C

If your worksheet is organized appropriately, you can use a frequency column when you create a graph.

You can use one frequency column for all graph variables, or a different one for each. Frequency variables are used in the order they are entered. Be sure they match the entry order of their corresponding graph variables.

Without a frequency column

C1
Colors
Red
Blue
Blue
Red
Red
Blue
Red

With a frequency column

C1	C2
Colors	**Frequency**
Red	4
Blue	3

GAPS: Session subcommand for displaying a gap in time on a graph

GAPS

Specifies a gap to be displayed for excluded data points. GAPS is used with the INCLUDE or EXCLUDE subcommand for graphs with time scales.

GAPS is a valid subcommand, with the use of INCLUDE or EXCLUDE for the following types of graphs:

- Area graphs
- Time series plots
- Control charts

GAPWIDTH and COFFSET: Session subcommands for the space between clusters and items in a cluster

Control the spacing between and within clusters on a multi-level categorical scale.

GAPWIDTH K

Use a value from 0 to 5 to specify the space between clusters. Using 0 minimizes the gap between the clusters, and 5 maximizes the gap.

COFFSET K

Use a value from –1 to 1 to specify the space between items in a cluster. A negative value overlap the items, using 0 places them next to one another, and 1 maximizes the space between items.

GRAPH: Session subcommand for controlling the graph region fill and border line

GRAPH [K K]

Controls the size, the fill type, and the fill color of the graph region, and the type, the color, and the thickness of the border line around the graph region.

The arguments represent the width and height of the graph region, respectively, and use inches. Values for arguments must be greater than or equal to 1 and less than or equal to 30. When used without arguments, the graph region maintains its default size.

Fill pattern and background fill color

Specifies the fill pattern and background fill color of the graph region.

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Graph region border

Specifies the type, color, and width of the border line.

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number. Enter one K for each bar or store the line sizes in C.

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph

MGRID K

Generates customizable horizontal or vertical lines that extend across the graph from the major and/or minor positions. You must customize major and minor, x-axis, y-axis and z-axis gridlines separately.

With MGRID K, the argument 1 places gridlines on minor ticks of the x-axis, 2 places them on the y-axis. Categorical scales do not have minor ticks and do not accept minor gridlines. 3 specifies the z-axis on 3D graphs.

SECSCALE

Use to place gridlines in relation to ticks on the secondary scale of the specified axis.

NOGRID

NOMGRID

If you've set graph preferences to display major or minor gridlines by default, use NOGRID and NOMGRID to remove gridlines on a graph-by-graph basis.

Type, color, and size of the gridlines

Specifies the type, color, and size of the gridlines that are invoked by the preceding command.

TYPE K

Specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256.

COLOR K

COLOR specifies the line color in K, which is a number from Numbers for colors to use in session commands on page 1255.

SIZE

Specifies the size of the gridlines.

GROUP: Session subcommand for specifying categorical variables for grouping

GROUP C [C...C]

Use GROUP when using two or more categorical variables with a dotplot, boxplot, interval plot, individual value plot, or bar chart.

Specify the first grouping variable with the graph command. Specify additional grouping variables as arguments to the GROUP subcommand.

GSAVE: Session subcommand for saving a graph in a file

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG

JPEG color

PNGB

PNG grayscale

PNGC

PNG color

TIFB

TIF grayscale

TIF

TIF color

BMPB

BMP grayscale

BMPC

BMP color

GIF

GIF

EMF

EMF

RESOLUTION K

Saves the graph at a resolution of K dots per inch.

HLABEL: Session subcommand for labeling histogram bars on a marginal plot with y-axis values

HLABEL

Labels histogram bars on a marginal plot with their y-axis values.

POSITION K "K"
Labels individual histogram bars. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

ENDPOSITION
After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*
TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

HLINE: Session subcommand for specifying attributes for historical stage lines

HLINE

Specify attributes for the historical stage lines on a control chart. The HLINE attributes apply to all historical stage lines on a chart.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number that is specified in Numbers for line types to use in session commands on page 1256 list. Enter one K for line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number that is specified in Numbers for colors to use in session commands on page 1255 list. Enter one K for each line or store the line colors in C.

SIZE K...K

SIZE C

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Label attributes

LSHOW K

Specifies whether to display the label.

Value of K	Label
0	Do not display the label
1 (default)	Display the label

TFONT *"font name"*
TFONT specifies the text font. Enclose the font name "K" in double quotation marks.

TCOLOR K
TCOLOR specifies the label color and uses integer arguments that are specified in Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Label positioning

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 specifies horizontal text.

ILABEL: Session subcommand for labeling individual values on a boxplot or interval plot

ILABEL [C]
Labels individual data values on a boxplot with text stored in C.

Labels individual data values on a boxplot or interval plot. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  Minitab matches the label column to the data column by worksheet row number. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the ROWNUM subcommand to label data values with their worksheet row number.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands affect only the individual label until ENDPOSITION is issued.

YVALUE
Labels each data value with its y-axis value.

ROWNUM
Labels each data value with its worksheet row number.

Labeling individual data values

POSITION K "K"
Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K
The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION
After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*
TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K
Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

INCLUDE and EXCLUDE: Session subcommands for including or excluding rows on a graph

INCLUDE

EXCLUDE

These commands are mutually exclusive: the last one issued is honored. Choose INCLUDE or EXCLUDE based on the degree of subsetting. For example, if you want to use all but a few select rows in creating your graph, it is more efficient to use EXCLUDE and name a small number of rows with the ROWS subsubcommand.

ROWS K...K

List all row numbers to be included or excluded with a space between each. Denote a patterned range by K:K/K. For example, 10:50/5 denotes all values from 10 to 50 by intervals of 5.

BRUSHED

Use to specify all brushed rows for subsetting.

WHERE *"expression"*

Specify a condition to be met for subsetting. This condition must be in double quotation marks. For example:

- If you want to subset where C1 is 2, enter `WHERE "C1 = 2"`.
- If you want C2 = a, enter `WHERE "C2 = ""a"""`.
- If you want C3 = 3/3/18, enter `WHERE "C3 = DATE(""3/3/18"")"`.

INDIVIDUAL: Session subcommand for displaying a symbol for each individual data value on a boxplot or an individual value plot

INDIVIDUAL [C...C]

Displays a symbol for each individual data value on a boxplot or an individual value plot.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main DOTPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

```
INDPLOT C1*C2;
  INDIVIDUAL C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns different attributes (color, type) to the symbols representing each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C

TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

INTBAR: Session subcommand for displaying a vertical line with horizontal lines at the endpoints of the confidence interval for the mean

INTBAR

Displays a vertical line with horizontal lines at the endpoints of the confidence interval for the mean.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main graph command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different bar attributes.

```
INDPLOT C1*C2;
  INTBAR C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns a different color to the interval bar representing each graph variable and displays a legend.

CIBAR [K]

Displays the confidence interval for the mean. The default is 95. Specify the confidence level K as a proportion (e.g. 0.95) or as a percent (e.g. 95). Enter a positive number between 0 and 100. CIBAR and ERRBAR are mutually exclusive, Minitab honors the last one entered.

ERRBAR

Displays standard error bars with K multiples of the standard error away from the mean of each subgroup. CIBAR and ERRBAR are mutually exclusive, Minitab honors the last one entered.

SIDE

Specifies how the error bars or confidence intervals are drawn.

Value of K	How the error bars or confidence intervals are drawn
−1	Lower one-sided
0	Two-sided
1	Upper one-sided

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies line type, which is a number from Numbers for line types to use in session commands on page 1256.Enter one K for each distribution fit or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each distribution fit or store the line colors in C.

SIZE K...K

SIZE C

SIZE specifies the line width, which can be any positive number. Enter one K for each distribution fit or store the line sizes in C.

Bonferroni confidence intervals

NOBONFERRONI (default)

NOBONFERRONI specifies that the confidence intervals are individual confidence intervals.

BONFERRONI

Specifies to use Bonferroni confidence intervals, which applies a family confidence level to the set of intervals.

Pooling the standard error

NOPOOL (default)

NOPOOL (default) specifies to not pool the standard error.

POOL

Pools standard error across all subgroups instead of calculating the standard error for each subgroup separately. If you display confidence intervals instead of standard error bars, the critical value from the t-distribution is calculated using the pooled degrees of freedom.

Positions where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, see Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

INTLAB: Session subcommand for labeling interval bar endpoints on an interval plot

INTLAB [C]

Labels interval bar endpoints on an interval plot. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  Minitab assigns labels from left to right across the graph. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands only affect the individual label until ENDPOSITION is issued.

YVALUE

Labels each data value with its y-axis value.

Labeling individual data values

POSITION K "K"... K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD

Specifies to display the label in bold font.

NOBOLD

Specifies to not display the label in bold font.

UNDERLINE

Specifies to underline the label.

NOUNDERLINE

Specifies to not underline the label.

ITALIC

Specifies to display the label in italic font.

NOITALIC

Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K

Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

IQRBOX: Session subcommand for displaying an interquartile range box on a boxplot

IQRBOX [C...C]

Displays an interquartile range box on a boxplot with the default box bottom at the 25th percentile and default box top at the 75th percentile.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different box attributes.

```
BOXPLOT C1*C2;
  IQRBOX C2.
```

VASSIGN

For ovelaid boxplots, assigns different attributes (color, fill type) to the interquartile range boxes representing each graph variable and displays a legend.

Box edge line

Specifies the type, color, and width of the box edge line.

ETYPE K...K

ETYPE C

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each bar or store the line types in C.

ECOLOR K...K

ECOLOR C

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

ESIZE K...K

ESIZE C

ESIZE specifies the line width, which can be any positive number. Enter one K for each bar or store the line sizes in C.

Box fill pattern and background fill color

Specifies the fill pattern and background fill color of the box.

TYPE K...K

TYPE C

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each bar or store the fill types in C.

COLOR K...K

COLOR C

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the fill colors in C.

Endpoints for the box

Determines the endpoints for the box. You can only use one QUARTILE, HINGE, or PERCENTILE subcommand. If you enter more than one, Minitab only uses the last subcommand.

QUARTILE

QUARTILE places the box endpoints at the first and third quartiles.

HINGE

HINGE places the box endpoints at the lower hinge and upper hinge.

Quartiles and hinges are both measures of the middle half of the data. The hinges may be a bit closer to the median than the quartiles are. In practice, you would not notice this difference. However, with rare data sets (usually small in size) the quartile method can cause the whiskers to be drawn into the box.

PERCENTILE K

PERCENTILE places the box endpoints at the specified percentile. Enter a number between 0 and 50.

Positions where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, see Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

JITTER and NOJITTER: Session subcommands for randomly offsetting data points to reveal overlapping points

JITTER [K K]

JITTER [K K K]

Randomly offsets data points to reveal overlapping points. The first K is the x-axis range and the second K is the y-axis range. For 3D plots, the third K is the z-range. The proportion constants control the maximum distance jittered. Enter a value between 0 and 1 for the x- and y-axis ranges, respectively.

BASE K

Sets the starting point for Minitab's random number generator. The generator uses this starting point until you set a new BASE or exit Minitab. To create a graph with the same jitter, use BASE with the same starting point K, provided the number of data values remains constant.

NOJITTER

Turns off JITTER.

LAYOUT and ENDLAYOUT: Session subcommands for specifying where a graph appears on a page

Specify where one or more graphs appear on one page. In LAYOUT mode, subsequent graphs that you create are displayed on a single page. After you create all the graphs you want, take Minitab out of LAYOUT mode and display the page by using ENDLAYOUT.

By default, Minitab places each graph at the same page coordinates, one on top of the other as if you were stacking transparencies, unless you specify different coordinates for each graph using the FIGURE subcommand.

You can annotate a page with titles, footnotes, text, symbols, lines, and polygons that are not specific to any graph, or make text charts and drawings without any graphs on them at all.

The following commands generate a single graph page with a bar chart in the top half and a time series graph in the bottom. The layout is titled "Sales Figures: 2015–2016."

```
LAYOUT;
 TITLE "Sales Figures: 2015-2016".
 CHART ('Sales') * 'Year';
 SUMMARIZED;
 BAR;
 FIGURE 0 1 0.5 1.
```

```
 TSPLOT 'Sales';
FIGURE 0 1 0 0.5.
```

LAYOUT
Begins a graph layout.

ENDLAYOUT
Ends a graph layout.

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Labels

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Regions

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LBRIGHT: Session subcommand for specifying the brightness of the lights that illuminate a surface plot

LBRIGHT K
Specifies the brightness (as an integer from 0 to 100) of the lights that illuminate the surface. LBRIGHT has no effect when WIREFRAME is used.

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph

LEGEND [K K K K]

Controls the size and placement of the legend within the figure region, the fill type and color of the legend, the type, color and thickness of the edge line around the legend, the font characteristics of the legend text, and the legend text displayed.

The arguments represent the x-minimum, x-maximum, y-minimum, and y-maximum values for the legend, respectively, and use figure units which range from 0 to 1. When used without arguments, the legend maintains its default size.

NOLEGEND

Use NOLEGEND to hide the legend.

Pattern and color of the legend background

Specifies the pattern and color of the legend background.

TYPE K

TYPE specifies the background fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background fill color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and size of the legend border

Specifies the type, color, and width of the legend border.

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number.

Font, color, size, and style of the legend body text

TFONT *"font name"*

TFONT specifies the font name for the text.

TCOLOR K

TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Font, color, size, and style of the legend header text

HFONT *"font name"*
TFONT specifies the font name for the text.

HCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

HTSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

HPSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

HBOLD
Specifies to display the text in bold font.

HNOBOLD
Specifies to not display the text in bold font.

HUNDERLINE
Specifies to underline the footnote.

HNOUNDERLINE
Specifies to not underline the text.

HITALIC
Specifies to display the text in italic font.

HNOITALIC
Specifies to not display the text in italic font.

Section number and section title

SECTION K
Specifies the section number.

STITLE *"string"*
Specify the section title.

CHEADER K *"string"*
Specify the header for column K.

BTEXT K K *"string"*
Specify the text for table row number K and column K.

STHIDE
Hide the section title.

CHHIDE
Hide the column headers.

RHIDE K...K
Hide row numbers K...K.

CHIDE K...K.
Hide table columns K...K.

LIGHT: Session subcommand for specifying the position, color, and visibility of the lights that illuminate a surface plot

LIGHT K K K K
Specifies the position, color, and visibility of the lights illuminating the surface.

The first argument for LIGHT specifies the number of the light (1, 2, or 3). The remaining arguments specify the position of the light in x-, y-, and z-coordinates. For example, `LIGHT 1 0 0 1` specifies that light 1 is to be centered on the x- and y-axes, one unit above the surface in the positive z-direction.

LIGHT has no effect when WIREFRAME is used.

COLOR K
COLOR specifies the color of the light in K, which is a number from Numbers for colors to use in session commands on page 1255.

VISIBILITY K
VISIBILITY specifies whether the light is on (0) or off (1).

LINE: The session subcommand for constructing a line from points on a graph

LINE K K...K K

LINE C C

Creates a line on a graph and controls attributes of the line.

You can also put lines anywhere on a graph by using graph editing. Using the graph editing is often easier, unless you want to repeatedly produce a graph with lines in the same positions, or you want to place a line in an exact location with predefined points. The REFERENCE on page 1146 subcommand is well suited to drawing horizontal and vertical reference lines and their labels in the data region. Use LINE for miscellaneous annotation lines.

To specify single lines, use the x y-coordinates of the line's start point for the first pair of arguments and the x y-coordinates of the line's end point for the second pair of arguments.

For a polyline, simply continue to add x y-coordinate pairs. For a polyline with many points, you can put x-arguments in one column, y-arguments in another and specify both columns.

To specify an x-axis coordinate that represents categories (as with CHART) in data units, specify the categories as axis positions 1, 2, ..., c (where c is the total number of categories) as they appear on the graph from left to right. You can specify points between categories. For example, a point midway between the third and fourth categories at location 44 on the y-axis would be 3.5 44.

LESTYLE K K

Specifies a starting and ending arrow style. The first argument specifies the line's starting arrow style, the second its ending arrow style. Use 0 for no arrow, 1 for a thin arrow, 2 for medium, and 3 for thick.

UNIT K

Specifies the unit system used by LINE and allows the use of data units in placing lines.

To place lines using figure units, use the argument 0. Use 1 to specify data units.

In cases of paneled graphs (including matrix plots), you may still place lines using data units if you also specify the panel with the UNIT command. UNIT 1 for example will place the line relative to the first panel of the graph and will read the LINE arguments in the data units of the first panel. You may choose to associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, LINE always uses graph units.

Type, color, and size of the line

Specifies the type, color, and size of the line.

TYPE K

TYPE specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256.

COLOR K

COLOR specifies the line color in K, which is a number from Numbers for colors to use in session commands on page 1255.

SIZE K

For SIZE, use a value that is greater than 0.

LOWESS: Session subcommand for fitting a LOWESS smoother to a scatterplot, a matrix plot, a histogram, or a time series plot

LOWESS [C...C]

Fits a lowess smoother to the data, where C...C can specify up to 3 grouping variables. When you use grouping variables, LOWESS generates a separately smoothed line for each group.

Fits a LOWESS smoother to a scatterplot, matrix plot, histogram, or time series plot. You can use a lowess smoother to explore the relationship between two variables without fitting a specific model. You can use LOWESS more than one time in a graph, each using different parameters.

Type, color, and size of the LOWESS line

Specifies the type, color, and size of the line.

TYPE K...K

TYPE C

TYPE specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each lowess smoother or store the line types in C.

COLOR K

COLOR C

COLOR specifies the line color in K, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each lowess smoother or store the line colors in C.

SIZE K

SIZE C

Specifies the line width, which can be any positive number. Enter one K for each lowess smoother or store the line sizes in C.

Algorithm

F K

Specifies the fraction K of the total number of points used to calculate the fitted values at each x-value. K can be any number from 0 to 1. The default is 0.5.

NSTEPS K

Specifies the number of iterations of smoothing used to limit the influence of outliers on the smoothed y-values. Each step reduces the weights given to outliers in the next iteration of weighted linear regression, based on the size of residuals in the previous lowess step. K can be any integer from 0 to 10. The default is 2. Some statisticians suggest that two robust steps adequately smooth outlier effects for most data.

Storage

Stores the fits and residuals for the lowess line. Not available for matrix plot or when the main command is in a LAYOUT.

FITS C...C

FITS stores the fitted values (new y-values) that comprise the lowess smoother.

RESIDUALS C...C

RESIDUALS stores the residual values (original y-values minus the fitted values).

If you specify one graph, specify a single column, even if you use groups with LOWESS. If you specify more than one graph, specify a residual column for each graph. The number of argument columns must be equal to the number of response columns.

You cannot store fits or residuals for matrix plot or when the graph is in a LAYOUT.

MARKER: Session subcommand for displaying a symbol at specified points on a graph

MARKER K K

MARKER C C

Specify the location of a marker.

MARKER displays a symbol (or marker) at the points specified in the subcommand.

You can also put markers anywhere on a graph using graph editing. Using the graph editing features is often easier, unless you want to repeatedly produce a graph with markers in the same positions, or you want to place markers in exact locations using predefined points.

UNIT K

Specifies the unit system used by MARKER and allows the use of data units in placing markers.

If you want to place markers using figure units, use the argument 0. Use 1 to specify data units.

With paneled graphs (including matrix plots), you can still place markers using data units if you also specify the panel with the UNIT command. For example, `UNIT 1` places the marker relative to the first panel of the graph and reads the MARKER arguments in the data units of the first panel. You can associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, MARKER always uses graph units.

Marker type, color, and size

Specifies the marker type, color, and size.

TYPE K

TYPE specifies marker type in K, which is a number from Numbers to use for symbols and markers in session commands on page 1257.

COLOR K

COLOR specifies the marker color, which is a number from Numbers for colors to use in session commands on page 1255.

SIZE K

For the marker size, choose a value greater than 0 and less than or equal to 15.

MEALAB: Session subcommand for labeling means on a boxplot or an interval plot

MEALAB [C]

Labels means on boxplot or interval plot. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  Minitab assigns labels from left to right across the graph. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands only affect the individual label until ENDPOSITION is issued.

YVALUE

Labels each data value with its y-axis value.

Labeling individual data values

POSITION K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K
Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

MEAN: Session subcommand for displaying symbol for each mean on a boxplot, interval plot, or individual value plot

MEAN [C...C]

Displays a symbol for each mean on a boxplot, interval plot, or individual value plot. The subcommands of MEAN allow you to control attributes (symbol type, color, size) for symbols (or groups of symbols) that represent means on a graph.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

```
BOXPLOT C1*C2;
  MEAN C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns a different color to the mean symbol that represents each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C

TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MEDIAN: Session subcommand for displaying a symbol for each median on a boxplot, interval plot, or individual value plot

MEDIAN [C...C]

Displays a symbol for each median on a boxplot, interval plot, or individual value plot. The subcommands of MEDIAN allow you to control attributes (symbol type, color, size) for symbols (or groups of symbols) that represent medians on a graph.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns a different color to the median symbol that represents each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C

TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MEDLAB: Session subcommand for labeling medians on a boxplot

MEDLAB [C]

Labels medians on a boxplot. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  Minitab assigns labels from left to right across the graph. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands only affect the individual label until ENDPOSITION is issued.

YVALUE

Labels each data value with its y-axis value.

Labeling individual data values

POSITION K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K
Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

NODOTFOOTNOTE: Session subcommand for suppressing footnotes on a dotplot

NODOTFOOTNOTE

Suppresses the default footnotes on a dotplot.

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph

NODTITLE

Suppresses default titles. NODTITLE does not work when you use LAYOUT.

NODSUBTITLE

Suppresses default subtitles.

NODFOOTNOTE

Suppresses default footnotes.

Note For an interval plot, use `NOINTFOOTNOTE` to suppress the footnote.

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs

NOEMPTY

Exclude empty values from attribute assignment rotation. Any combination of group levels that do not contain data are not assigned attributes. By default, groups that have empty values are included in attribute assignment rotation.

NOMISS

Exclude missing groups from the graph that is being created. If there is a group with a missing level for any group combination, then that row is excluded from the graph. By default, groups with missing levels are included in the graph.

NOFRAME: Session subcommand for suppressing lines and labels on a graph

NOFRAME

Specifies to suppress axis lines and labels, tick lines and labels, reference lines and labels, and grid lines in a graph.

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph

NOHLEGEND

Use NOHLEGEND to suppress display of the legend of hold values.

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote

NOPERFOOTNOTE

Suppresses the personal footnote stamped on a graph when set up in **File** > **Options** > **Graphics** > **Annotation** > **My Footnote**.

NOSEPSUBTITLE: Session subcommand for subtitles on separate graphs

NOSEPSUBTITLE

Suppresses default subtitles on separate graphs when you make multiple graphs.

OLAB: Session subcommand for labeling outliers on a boxplot

OLAB [C]

Labels outliers on a boxplot. Minitab determines the label by the main command line argument or a subcommand.

- Enter a column on the main command line to use labels that are stored in a column. If no argument is given on the main command line, Minitab labels the points with their y-axis values. The label column can be numeric, text, or date/time. Labels can be any combination of up to 80 characters, including spaces.

  Minitab assigns labels from left to right across the graph. If the label column contains more values than are needed, Minitab ignores the extra labels. If the label column contains less values than are needed, the remaining data values are not labeled.

- Use the YVALUE subcommand to label data values with their y-axis value.
- Use the ROWNUM subcommand to label data values with their worksheet row number.
- Use the POSITION subcommand to label an individual data value. After issuing POSITION, the font and positioning subcommands only affect the individual label until ENDPOSITION is issued.

YVALUE

Labels each data value with its y-axis value.

ROWNUM

Labels each data value with its worksheet row number.

Labeling individual data values

POSITION K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD

Specifies to display the label in bold font.

NOBOLD

Specifies to not display the label in bold font.

UNDERLINE

Specifies to underline the label.

NOUNDERLINE

Specifies to not underline the label.

ITALIC

Specifies to display the label in italic font.

NOITALIC

Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

PLACEMENT K K

Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

OUTLIER: Session subcommand for displaying a symbol for each outlier on a boxplot

MEAN [C...C]

Displays a symbol for each outlier on a boxplot, where C...C can specify up to 4 grouping variables, and restores outlier symbols when another data display command turns them off (for example, if you type IQRBOX, OUTLIER is turned off).

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

```
BOXPLOT C1*C2;
  OUTLIER C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns a different color to the mean symbol that represents each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C

TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph

OVERLAY

When you create multiple graphs with the same command, use OVERLAY to specify that they display on the same data region.

OVERLAY displays the graph variables on the same graph and assigns different attributes to each. A legend is produced to link the different data representations to the graph variables.

When overlaying charts, dotplots, interval plots, individual value plots, or boxplots, OVERLAY causes the graph variables to be treated as groups on a single scale level.

PANEL: Session subcommand for paneling graphs

PANEL [C...C]

PANEL is used in several ways. Use it to panel multiple graphs together, or to divide a single graph into panels defined by a paneling variable. You can also combine these methods and panel both by graph variables and groups.

Used without an argument, PANEL specifies the placement of multiple graphs created with a single graph command. Arguments to PANEL are By variables that split the graph.

SPLIT K

Split the data of a time series plot or area graph into panels. Available only with time series plot and area graph when a group variable is not entered in the PANEL subcommand. SPLIT separates sets of K data into panels. This subcommand is especially useful with vary large data sets or cycling data. SPLIT is available as a subcommand to control charts.

NOALTERNATE

By default, Minitab places the ticks of adjacent panels on alternating sides of the graph. Invoking NOALTERNATE places ticks on the same side of adjacent panels.

Panel order

Use one of these subcommands when you have paneled both by model and by a paneling variable.

VFIRST

Use one of these subcommands when you have paneled both by model and by a paneling variable. VFIRST arranges panels by model first. Thus, when read left to right each row displays one model with a varying paneling variable.

VLAST

VLAST reverses this arrangement.

Row and column configuration

Specify the row and column configuration as well as the margins separating them.

RC K K

RC takes two arguments to customize the panel layout. The first specifies the number of rows in the layout, the second specifies the number of columns.

MARGIN K

The optional MARGIN specifies the margin size separating all panels. Use a value from 0 to 0.25.

Label type, font, style, color, and size

Specify the label type, font, style, color, and size.LABEL

TYPE K [K]

Specify the degree of information to include in the panel labels. The first argument to TYPE refers to Y–X information, the second to group information.

Value of the first K	Degree of Y–X information to include in the panel labels
0	No Y–X information
1	Y and X information
2	Y information
3	X information

Value of the second K	Degree of group variable information to include in the panel labels
0	No group variable information
1	Group level information

Value of the second K	Degree of group variable information to include in the panel labels
2	Group variable and level information

TFONT *"font name"*
TFONT specifies the font name for the labels.

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

POLYGON: Session subcommand for constructing a polygon from points on a graph

POLYGON K K...K K

POLYGON C C
POLYGON constructs a polygon from the points specified in the subcommand. Each pair of arguments defines a vertex of the polygon. The first argument of a pair is the x-coordinate, the second, the y-coordinate. The first and last points connect to close the polygon.

For a polygon with many points, you can put x-arguments in one column, y-arguments in another and specify both columns.

To specify an x-axis coordinate when the x-axis represents categories (as with CHART), specify the categories as axis positions 1, 2, ..., c (where c is the total number of categories) as they appear on the graph from left to right. You can specify points between categories. For example, a point midway between the third and fourth categories at location 44 on the y-axis would be 3.5 44.

You can also put polygons anywhere on a graph using the graph editing features. Using the graph editing features is often easier, unless you want to repeatedly produce a graph with polygons in the same positions, or you want to place a polygon in an exact location using predefined points.

When POLYGON is a direct subcommand of LAYOUT on page 1116, use page units to specify the argument pairs.

UNIT K

To place polygons using figure units, use the argument 0. Use 1 to specify data units.

With paneled graphs (including matrix plots), you can still place polygons using data units if you also specify the panel with the UNIT command. For example, `UNIT 1` places the polygon relative to the first panel of the graph and reads the POLYGON arguments in the data units of the first panel. You can associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, POLYGON always uses graph units.

Fill pattern and background fill color

Specifies the fill pattern and background fill color of the polygon.

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the line

Specifies the type, color, and size of the polygon's edge.

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number. Enter a value that is greater than 0.

PROJECT: Session subcommand for extending projection lines from each point to the x-axis

PROJECT [C...C]

Extends projection lines from each point down to the horizontal, or x-axis.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.

- Grouping columns must be categorical variables that are included on the main graph command line, or with the GROUP subcommand. In the following example, C3 is the grouping column, and, for each unique value in C3, Minitab assigns different line attributes and symbol attributes.

```
PLOT C1*C2;
  PROJECT C3.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN

For overlaid graphs, assigns a different color to the mean symbol that represents each graph variable and displays a legend.

DIRECTION

Specifies the direction the project lines extends away from the data values. DIRECTION does not apply to HISTOGRAM.

Value of K	Direction of the lines
1	Extends horizontally from the data values to the base (the yz-plane for 3D graphs)
2	Extends from the data values to the base (the xz-plane for 3D graphs) (default for 2D graphs)
3	Extends vertically from the data values to the base (the xy-plane for 3D graphs)

BASE

Specifies the for the project lines. For more information, go to Base position for project lines, area, and bar on page 1217. Project lines extend from the data value to the base. The default base is the minimum data value for the y-scale.

Enter one K for each base or store the bases in C. If one value is specified it applies to all points.

You can draw a base line by using REFERENCE on page 1146.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each line or store the line types in C.

SIZE K

SIZE specifies the line width. Enter one K for each line or store the line types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

PROPORTIONAL and NOPROPORTIONAL: Session subcommands for making the boxes on a boxplot proportional to the square root of the number of observations in the boxes

PROPORTIONAL

Makes the width of each box on a boxplot proportional to the square root of the number of observations in the box.

NOPROPORTIONAL (default)

Returns the box to the standard width.

PTILES: Session subcommand for specifying the location of the percentile lines on a graph

PTILES K [K...K]

PTILES generates reference lines on percentile scales. A percentile line runs from the percentile scale to the distribution fit, and then to its corresponding point on the data value scale. The points on both scales are labeled.

A percentile line is placed for each argument given. Arguments are in percentiles.

DVALUE

Specifies that the argument to the previous PTILES subcommand are not percentiles, but data values.

PSHOW K

Specifies the percentile line label display.

Value of K	Percentile lines that are displayed
0	Hide labels
1 (default)	Show both labels
2	Show x-scale label
3	Show y-scale label

MODEL K
Use only when you create multiple graphs with a single command. Allows you to apply its related subcommand and subsubcommands to a specific graph. The argument refers to the order the graphs are created. MODEL 1 affects the first graph created, MODEL 2 the second graph, and so on.

Attributes for all labels

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Attributes for percentile line labels

PLABEL [K]
Customize the label corresponding to the PTILE argument. Use any subcommands (TFONT through PLACEMENT) to customize the percentile line label, then issue ENDLABEL.

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

BOLD

Specifies to display the text in bold font.

NOBOLD

Specifies to not display the text in bold font.

UNDERLINE

Specifies to underline the text.

NOUNDERLINE

Specifies to not underline the text.

ITALIC

Specifies to display the text in italic font.

NOITALIC

Specifies to not display the text in italic font.

PLACEMENT K K

Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

PPOSITION K and ENPPOSITION

Specifies the panel position for customizing a particular percentile line label or information label. PPOSITION 3 would refer to the third panel.

When customizing a percentile line label, use ANGLE, OFFSET, or PLACEMENT subcommands by panel position to modify the label then issue ENPPOSITION.

ENDLABEL
Specifies the end of the PLABEL command.

Attributes for information labels

ILABEL
Specifies the value (information label) calculated by Minitab given the PTILE value.

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

STRING

Specify text (in quotation marks if a supplied constant) for a particular information label, for example, "50th percentile."

PPOSITION K and ENPPOSITION

Specifies the panel position for customizing a particular percentile line label or information label. PPOSITION 3 would refer to the third panel.

When customizing a percentile line label, use ANGLE, OFFSET, or PLACEMENT subcommands by panel position to modify the label then issue ENPPOSITION.

GPOSITION K and ENGPOSITION

Specifies the panel position for customizing a particular information label. GPOSITION 3 would refer to the third group. Use any subcommands (TFONT through STRING) by panel position to modify the information label then issue ENGPOSITION.

ENDLABEL

Specifies the end of the ILABEL command.

RBOX: Session subcommand for displaying a range box on a boxplot

RBOX [C...C]

Displays a range box on a boxplot that extends from the minimum to the maximum data value, where C...C can specify up to 4 grouping variables.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- Grouping columns must be categorical variables that are included on the main BOXPLOT command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different box attributes.

```
BOXPLOT C1*C2;
    RBOX C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN K...K

VASSIGN C

For overlaid boxplots, assigns different attributes (color, fill type) to the range boxes representing each graph variable and displays a legend.

Fill pattern and background fill color of the box

Specifies the fill pattern and background fill color of the box.

TYPE K...K

TYPE C

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256. Enter one K for each bar or store the fill types in C.

COLOR K...K

COLOR C

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the fill colors in C.

Type, color, and width of box edge line

ETYPE K...K

ETYPE C

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each bar or store the line types in C.

ECOLOR K...K

ECOLOR C

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each bar or store the line colors in C.

ESIZE K...K

ESIZE C

ESIZE specifies the line width, which can be any positive number. Enter one K for each bar or store the line sizes in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph

RECTANGLE K K K K [K K K K]

RECTANGLE displays a rectangle at the points specified in the subcommand.

You can also put a rectangle anywhere on a graph using graph editing. Using the graph editing features is often easier, unless you want to place a rectangle in exact locations using predefined points.

A rectangle is defined by either four or eight arguments. Four arguments are used to represent the lower left and upper right coordinates (X Y X Y) when the rectangle is a square. This is sufficient for squares that are not rotated or are rotated exactly 90 degrees. If eight parameters are entered they represent X Y coordinates (lower left, lower right, upper right, upper left) for the four corners of the rectangle.

UNIT K

To place rectangles using figure units, use the argument 0. Use 1 to specify data units.

With paneled graphs (including matrix plots), you can still place rectangles using data units if you also specify the panel with the UNIT command. For example, `UNIT 1` places the rectangle relative to the first panel of the graph and reads the RECTANGLE arguments in the data units of the first panel. You can associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, RECTANGLE always uses graph units.

Fill pattern and background fill color

Specifies the fill pattern and background fill color of the rectangle.

TYPE K

TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the line

Specifies the type, color, and size of the rectangle's edge.

ETYPE K

ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K

ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K

ESIZE specifies the line width, which can be any positive number. Enter a value that is greater than 0.

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph

REFERENCE K [K...K]

REFERENCE K C

REFERENCE generates horizontal and vertical reference lines and their labels.

The first argument specifies the axis (1 = x-axis, 2 = y-axis, 3 = z-axis) at which the REFERENCE lines start/end. The arguments that follow give the locations of the reference lines in data units (K) or a column (C) that contains the locations of the reference lines in data units. If you specify a scale transformation, be sure to use the new scale units when placing reference lines. With a log transformation, reference line locations are in transformed units; with a power transformation, transformed units.

Label attributes

SECSCALE

Places the reference line in reference to the secondary scale of the stated axis.

SIDE K

Specifies the side where the reference line label is placed.

LABEL *"label"..."label"*

LABEL C

Specifies custom reference line labels or a column where the labels are located.

TFONT K

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

BOLD

Specifies to display the label in bold font.

NOBOLD

Specifies to not display the label in bold font.

UNDERLINE

Specifies to underline the label.

NOUNDERLINE

Specifies to not underline the label.

ITALIC

Specifies to display the label in italic font.

NOITALIC

Specifies to not display the label in italic font.

PLACEMENT K K

Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K

Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K

Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

Label positioning

MODEL K

Use MODEL only when you create multiple graphs with a single command. Allows you to apply its related subcommand and subsubcommands to a specific graph. The argument refers to the order the graphs are created. MODEL 1 affects the first graph created, MODEL 2 the second graph, and so on.

PPOSITION and ENPPOSITION

Specifies the panel position for customizing a particular reference line label. PPOSITION 3 would refer to the third panel label. Use ANGLE, OFFSET, or PLACEMENT subcommands to modify the label then issue ENPPOSITION.

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each line or store the line types in C.

SIZE K

SIZE specifies the line width. Enter one K for each line or store the line types in C.

REGRESS: Session subcommand for fitting a regression line to data on a graph

REGRESS [C...C]

Fits a regression line to the data, where C...C can specify up to 3 grouping variables. If you use grouping variables, REGRESS generates a separately smoothed line for each group.

Use to fit a least squares regression line to a scatterplot or matrix plot. Use the regression line to examine the relationship between the response variable (y) and the predictor variable (x). If you hover the mouse pointer over the fitted line, Minitab displays the equation. You can use REGRESS more than once in a graph, each with a different polynomial order.

NOCONSTANT

Omits the b_o term (the intercept) from the equation. For example, for the default linear model, Minitab fits:

$Y = \beta_1 X_1 + e$

RPOLY K

Specifies the order of the regression model.

Value of K	Order of the regression model
1	Fits a linear model
2	Fits a quadratic model
3	Fits a cubic model

Line type, color, and size

Specifies the line type, color, and size.

TYPE K...K

TYPE C

TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256. Enter one K for each line or store the line types in C.

COLOR K...K

COLOR C

COLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each line or store the line types in C.

SIZE K...K

SIZE C

SIZE specifies the line width, which can be any positive number. Enter one K for each regression fit or store the line sizes in C.

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs

Specify that one or more axes are the same when creating multiple graphs.

`SAME 1` makes the x-scale the same for all graphs. `SAME 2` makes the y-scale the same. `SAME 2` for dotplot means the same number of data values per dot. `SAME 1 2` makes the x- and y-scales the same.

SAME does not work for graphs with categorical or time scales.

SCALE: Session subcommand for customizing the axes and ticks of a graph

SCALE K

Use the argument 1 to denote the x-axis, 2 for the y-axis, 3 for the z-axis. All subcommands that follow will affect this scale.

LDISPLAY K K K K

LDISPLAY refers to the low sides of the scales (bottom and left side). The four arguments refer to the axis line, major ticks, major tick labels, and minor ticks, respectively. For categorical and time scales, the fourth argument is ignored. The argument 1 displays the corresponding item, 0 hides it.

HDISPLAY K K K K

HDISPLAY refers to the high (top and right) sides. The four arguments refer to the axis line, major ticks, major tick labels, and minor ticks, respectively. For categorical and time scales, the fourth argument is ignored. The argument 1 displays the corresponding item, 0 hides it.

SECSCALE C...C

Displays a secondary scale when you have multiple x or y variables on the same graph. Useful when overlaying graphs with dissimilar scales. In the following example, C1 is the primary y-variable and C3 is the secondary y-variable. Available only with the continuous scales of overlaid scatterplots and time series plots.

```
PLOT C1*C2 C3*C2;
 SCALE 2;
  SECSCALE C3;
 OVERLAY.
```

MIN K

Set the minimum value for the scale. For continuous scales only. This subsubcommand replaces the subcommand MINIMUM. Not available with categorical scales.

MAX K

Set the maximum value for the scale. For continuous scales only. This subsubcommand replaces the subcommand MAXIMUM. Not available with categorical scales.

Type, color, and size of the axis lines and ticks

TYPE K
TYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

COLOR K
COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

SIZE K
SIZE specifies the size of the axis lines and ticks.

Continuous scale ticks

TICK K...K

TICK C
Set the positions for tick marks. You can list the positions individually, set a patterned range, or name a column that contains the tick positions. The patterned range 10:50/5, for example, places ticks at every five units from 10 to 50. You cannot use both TICK and NMAJOR.

NMAJOR K
Set the number of major ticks. Minitab positions the ticks.

NMINOR K
Set the number of minor ticks between each pair of major ticks.

ORIENT K
Set the ticks to point out (K=1), in (K=2), or out and in (K=3). Not available with 3D graphs.

LENGTH K
Specify the length of the tick lines, from 0 to 0.05. Not available with 3D graphs.

Categorical and time scale ticks

TLEVEL K
Use TLEVEL to specify the units used to place ticks. The argument 1 refers to the outermost scale level, 2 to the second outermost, and so on. TLEVEL is only available with time scales.

TSTART K
Use TSTART to specify the starting point on the given level when used in conjunction with TLEVEL, and on the innermost level in all other cases. The argument 2 to refer to the month scale gives a starting point of February. When referring to a non-time variable, the starting point is the second value on the innermost scale.

TINCREMENT K
Use TINCREMENT to specify the frequency of the ticks. The argument 2 places ticks at every other value on the innermost level or the one set with TLEVEL.

Tick labels

LABEL *"string"..."string"*

LABEL C

Specify labels for major ticks, either by listing them individually or by naming a column that contains the labels. For categorical and time scales you need to list a single label only once and Minitab will cycle that label.

LEVEL K

Using the subsubcommand LEVEL K applies the named labels to a particular variable level on a multi- level scale, such as a nested boxplot. The label closest to the axis is referenced by an argument value of 1, the second closest by 2, and so on.

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K

TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K

TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K

PSIZE specifies the size of the text in points. K can be any positive number.

ANGLE K

Specify the angle of the tick labels in degrees from the horizontal.

BOLD

Specifies to display the label in bold font.

NOBOLD

Specifies to not display the label in bold font.

UNDERLINE

Specifies to underline the label.

NOUNDERLINE

Specifies to not underline the label.

ITALIC

Specifies to display the label in italic font.

NOITALIC

Specifies to not display the label in italic font.

TSHOW K [K [K [K]]]

Specify the tick labels displayed, by level. The level farthest from the axis is referenced by the argument 1, the second farthest by 2, and so on. Tick labels are shown for all levels by default.

Transformations

Transform the scale. Unless you have negative values you cannot use negative arguments.

LOGS K
Use LOGS K (where K is an integer from 2 to 10) to specify the base of the log transformation.

LOGE
LOGE for log base e.

POWER K
POWER K (where K is an integer from −5 to 5) to specify the degree of power.

STANDARD
Use standardized units for the y-scale. You can only use this subcommand when you specify SCALE2

Model reference

When a set of SCALE subsubcommands refer to a specific model of an overlaid graph, the commands should be bound with MODEL K and ENDMODEL K, where K refers to the model to be modified. Within SCALE, MODEL is valid with MIN/MAX, TICK, NMAJOR, NMINOR, and LABEL.

MODEL K
Specifies the model.

ENDMODEL K
Specifies the end of the model.

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable

SEPARATE [C...C]
When you use a paneling variable, SEPARATE places the different groups in separate graphs. Arguments to SEPARATE are By variables that split the graph into panels with one panel for each level or level combination of the By variable.

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot

SFIT
Controls the attributes of the surface on a surface plot.

COLOR K
COLOR specifies the color of the surface in K, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K
ESIZE specifies the wireframe width in K, which can be any positive number.

ECOLOR K
ECOLOR specifies the color of the wireframe in K, which is a number from Numbers for colors to use in session commands on page 1255.

SLABEL: Session subcommand for labeling the slices on a pie chart

SLABEL

Labels slices on a pie chart. If you do not use any subcommand, Minitab labels the slices with their category name.

Label attributes and positioning subcommands can apply to all labels (as subcommands of SLABEL) or to specified labels (as subcommands of POSITION).

PCATEGORY

Labels pie slices with their category names.

PFREQUENCY

Labels pie slices with their frequency.

PERCENT

Labels pie slices with their percentage of the whole.

PLINE

PLINE draws a line from the label to the slice.

TYPE K

TYPE specifies the line type in K, which is a number from Numbers for line types to use in session commands on page 1256.

COLOR K

COLOR specifies the line color in K, which is a number fromNumbers for colors to use in session commands on page 1255.

SIZE K

SIZE specifies the line width, which can be any positive number. Enter one K for each line or store the line sizes in C.

Labeling individual data values

POSITION K "K"

Labels individual data values. The first K is the row number; the second "K" is the label. The attributes and positioning subcommands can be used as subcommands to POSITION. You can use POSITION more than one time.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all labels. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

Label font, color, and size

TFONT *"font name"*

TFONT specifies the font name for the labels.

TCOLOR K
TCOLOR specifies the label color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the label

BOLD
Specifies to display the label in bold font.

NOBOLD
Specifies to not display the label in bold font.

UNDERLINE
Specifies to underline the label.

NOUNDERLINE
Specifies to not underline the label.

ITALIC
Specifies to display the label in italic font.

NOITALIC
Specifies to not display the label in italic font.

Positioning

Specifies the placement of the label in relationship to the anchor position.

OFFSET K K
Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

SLICE: Session subcommand for specifying the attributes of slices on a pie chart

SLICE
Specifies the attributes of slices on a pie chart.

POSITION K...K
Lists the positions where subsequent subcommands apply. Any slice subcommand can be used with POSITION.

ENDPOSITION
After ENDPOSITION is executed, additional subcommands apply to all slices.

Type and color of the slices

TYPE K
TYPE specifies the fill pattern, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K
COLOR specifies the background color, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and size of the slice edge lines

ETYPE K
ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K
ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K
ESIZE specifies the line width, which can be any positive number. Enter a value that is greater than 0.

Separating the slices

EXPLODE
EXPLODE separates the slices from one another by moving them out radially from the center by LENGTH K.

LENGTH K
Enter a value between 0 and 1 for the LENGTH argument.

SOFFSET: Session subcommand for offsetting points from the center

SOFFSET
Offsets identical points symmetrically from the center.

STAMP: Session subcommand for specifying columns that contain time values for observations

STAMP C [C...C]
Specify up to three columns that contain the time values for the observations.

SUBTITLE: Session subcommand for adding a subtitle to a graph

SUBTITLE *["subtitle"]*
Adds a subtitle to a graph. Using SUBTITLE with no argument displays the default subtitle. Using SUBTITLE one time with an argument overwrites the default subtitle.

Font, color, and size of the subtitle

TFONT *"font name"*
TFONT specifies the text font.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the subtitle

BOLD
Specifies to display the subtitle in bold font.

NOBOLD
Specifies to not display the subtitle in bold font.

UNDERLINE
Specifies to underline the subtitle.

NOUNDERLINE
Specifies to not underline the subtitle.

ITALIC
Specifies to display the subtitle in italic font.

NOITALIC
Specifies to not display the subtitle in italic font.

Position of the subtitle

LEFT
Positions the subtitle on the left of the figure region.

CENTER (default)
Positions the subtitle at the center of the figure region.

RIGHT
Positions the subtitle on the right of the figure region.

OFFSET K K
Moves a subtitle away from the position specified with LEFT, CENTER or RIGHT. The first argument specifies how far to move the subtitle horizontally (in figure units). The second argument specifies the how far to move the subtitle vertically (in figure units). The default offset is 0.0 0.0.

SYMBOL: Session subcommand for displaying a symbol for each data value

SYMBOL [C...C]
Displays a symbol for each data value, where C...C can specify up to 4 grouping variables for CHART and 3 for other graphs.

For grouping variables specified on the main command line:

- Can use up to four grouping columns.
- CHART can use up to four grouping columns. Grouping columns must be categorical variables that are included on the main CHART command line, or with the GROUP subcommand. In the following example, C2 is the grouping column, and, for each unique value in C2, Minitab assigns different symbol attributes.

```
CHART C1;
  SYMBOL C2.
```

Note For an overview and examples, go to Using groups in graphs on page 1265. For more information, go to Using the SYMBOL and COLOR subcommands for PLOT on page 1266.

VASSIGN
For overlaid dotplots, assigns different attributes (color, type) to the dots representing each graph variable and displays a legend.

Symbol type, color, and size

Specifies the symbol type, color, and size. You can enter a single K to apply to all symbols. If you use group variables, each K applies to a single group.

TYPE K...K

TYPE C
TYPE specifies the symbol type, which is a number from Numbers to use for symbols and markers in session commands on page 1257. Enter one K for each symbol or store the symbol types in C.

COLOR K...K

COLOR C

COLOR specifies the symbol color, which is a number from Numbers for colors to use in session commands on page 1255. Enter one K for each symbol or store the symbol colors in C.

SIZE K

SIZE specifies the size of the symbol. Enter one K for each symbol or store the symbol types in C.

Position where subsequent attribute subcommands apply

POSITION K...K

Lists the positions where subsequent attribute subcommands apply. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

MODEL K

The MODEL subcommand allows you to use POSITION with paneled graphs. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

ENDPOSITION

After ENDPOSITION is executed, additional subcommands apply to all models. For an example, go to Using the POSITION and MODEL subcommands on page 1268.

TABLE and NOTABLE: Session subcommand for controlling the table within the figure region

TABLE

Controls the size and the placement of the table within the figure region, the fill type and color of the legend, the type, color and thickness of the edge line around the table, the font characteristics of the table text, and the table text that is displayed.

NOTABLE

Hides the table.

Fill pattern and background fill color of the table

Specifies the fill pattern and background fill color of the table.

TYPE K

TYPE specifies the fill pattern in K, which is a number from Numbers for fill types to use in session commands on page 1256.

COLOR K

COLOR specifies the background color in K, which is a number from Numbers for colors to use in session commands on page 1255.

Type, color, and width of the table border

ETYPE K
ETYPE specifies the line type, which is a number from Numbers for line types to use in session commands on page 1256.

ECOLOR K
ECOLOR specifies the line color, which is a number from Numbers for colors to use in session commands on page 1255.

ESIZE K
ESIZE specifies the line width, which can be any positive number.

Font, color, and size of the table text

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

Style of the table text

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Font, color, and size of the header text

HFONT *"font name"*
HFONT specifies the font name for the text.

HCOLOR K
HCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

HTSIZE K
HTSIZE specifies the size of the text in Minitab units. K can be any positive number.

HPSIZE K
HTSIZE specifies the size of the text in Minitab units. K can be any positive number.

Style of the header text

HBOLD
Specifies to display the header text in bold font.

HNOBOLD
Specifies to not display the header text in bold font.

HUNDERLINE
Specifies to underline the header text.

HNOUNDERLINE
Specifies to not underline the header text.

HITALIC
Specifies to display the header text in italic font.

HNOITALIC
Specifies to not display the header text in italic font.

Section

SECTION K
Specify the section number.

STITLE *"string"*
Specify the section title.

CHEADER K *"string"*
Specify the header for table column K.

BTEXT K K *"string"*
Specify the text for table row number K and column K.

STHIDE
Hide the section title.

CHHIDE
Hide the column headers.

RHIDE K...K
Hide table row numbers K...K.

CHIDE K...K
Hide table columns K...K.

TEXT: Session subcommand for displaying text on a graph

TEXT K K *"text"*

TEXT C C C
Displays a text string anywhere on a graph. Use the x coordinate as the first argument and the y coordinate as the second argument. Enclose the text string in quotation marks.

To label the points on a graph, you can also use DATLAB on page 1089 with many graph commands.

When placing many text strings you can put x-coordinates in one column, the corresponding y-coordinates in a second, and the corresponding text strings in a third. Use these columns in this order as arguments to TEXT.

To specify an x-axis coordinate that represents categories (as with CHART) in data units, specify the categories as axis positions 1, 2, ..., c (where c is the total number of categories) as they appear on the graph from left to right. You can specify points between categories. For example, a point midway between the third and fourth categories at location 44 on the y-axis would be 3.5 44.

BOX K K K K
Specify the coordinates of the text box that will contain the text. The four arguments represent the x-minimum, x-maximum, y-minimum, and y-maximum coordinates of the text box

UNIT K
Specifies the unit system used by TEXT and allows the use of data units in placing text on paneled graphs.

To place text using figure units, use the argument 0. Use 1 to specify data units.

With paneled graphs (including matrix plots), you can still place text using data units if you also specify the panel with the UNIT command. For example, `UNIT 1` places the text in the first panel of the graph and reads the TEXT arguments in the data units relative to the first panel. You can associate the placement of annotation with data units or figure units to accommodate changes in scales or graph size. As a subcommand to LAYOUT, TEXT always uses graph units.

Attributes for the text

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Positioning of the text

PLACEMENT K K
Specifies where the label attaches to the anchor position. The first argument specifies the horizontal placement. The second argument specifies the vertical placement.

Vertical placement (second K argument)	Horizontal placement (first K argument)		
	−1	0	1
1			
0			
−1			

OFFSET K K
Moves a label away from the anchor position specified with PLACEMENT. The first argument specifies how far to move the label horizontally (in figure units). The second argument specifies the how far to move the label vertically (in figure units). The default offset is 0.0 0.0.

ANGLE K
Rotates the text counterclockwise by K degrees. K can be any number between −360 and 360. The default value of 0 produces horizontal text. You may need to change the POSITION and OFFSET values to achieve certain angle effects.

TITLE: Session subcommand for adding a title to a graph

TITLE *["title"]*
Adds a title to a graph.

- You can use TITLE more than one time. Each additional title goes below the preceding title.
- Using TITLE with no argument displays the default title. Using TITLE with an argument replaces the default title.
- Default titles precede additional titles.

TITLE works the same for the basic graph commands and control chart commands. You can suppress the default title with the subcommand NODTITLE.

Attributes for the title text

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Positioning

LEFT
Puts the title in the left of the figure region.

CENTER (default)
Puts the title in the center of the figure region.

RIGHT
Puts the title in the right of the figure region.

OFFSET K K

Moves a title away from the position specified with LEFT, CENTER or RIGHT. The first argument specifies how far to move the footnote horizontally (in figure units). The second argument specifies the how far to move the footnote vertically (in figure units). The default offset is 0.0 0.0.

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph

TRANSPOSE

TRANSPOSE interchanges the position of the x and y-axis on a graph. Transposing does not change the variable or name of the axis: the x-axis and y-axis remain the same whether they are horizontal or vertical. Keep this in mind when you use graphics options that require you to specify locations or directions in terms of the axis, such as gridlines or reference lines.

When TRANSPOSE is invoked, the data region is in effect turned on its side; the y-scale and x-scale trade places.

NOTRANSPOSE

Use to specify that a particular graph model is not to be transposed. Useful if you have specified (through preferences) that probability plots are to be transposed by default.

TSHOW: Session subcommand for specifying the level of tick labels that are displayed

TSHOW K [K [K [K]]]

Specify the tick labels displayed, by level. The level farthest from the axis is referenced by the argument 1, the second farthest by 2, and so on. Tick labels are shown for all levels by default.

VASPECT: Session subcommand for specifying the aspect ratio of the data box

VASPECT K K K

Specifies the aspect ratio of the data box in x, y, z order. Without arguments, VASPECT specifies that the ratio be proportional to the data range in each direction.

Some common reasons for changing aspect ratios are as follows:

- To make similar units be the same length on different axes. For example, suppose you have a range of 50 cm for x and 30 cm for y, and you want one centimeter to be represented by the same length on each axis. You can use VASPECT without arguments, or you can use VASPECT with arguments of 50 for x, 30 for y, and whatever makes the graph look good for the z.
- To make the graph a certain shape. For example, the default aspect ratio of 3:3:2 is a pleasing aspect ratio often known as a golden triangle. By making the graph a different shape, you can also achieve effects such as emphasizing changes in one variable (say, a variable with a large range) over another.

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph

VBOX K

Specifies the display of the box surrounding the data. VBOX does not control the display of axis lines, which are separate from the box and displayed behind.

Value of K	Display of the box faces
0	Hide all faces.
1 (default)	Show the back faces.
2	Show all the faces.

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units

VFIELD K K

Specifies the width and height of the field of view in object units. You can best think of object units in terms of the cube that displays three-dimensional graphs. The center of the cube is the origin (0, 0, 0) in object unit terms. Each edge of the cube is 2 object units long. This makes the shortest distance to an edge from the center of the cube 1 object unit. Larger values create a larger field of view and result in a smaller data box. The default is 3.75, with values between 2.5 and 4.0 being optimal for most applications.

VPOSITION: Session subcommand for specifying the view position as a ratio

VPOSITION K K K

Specifies the view position as a ratio in x, y, and z. The view position is the point from which you are viewing the data box. For example, at `VPOSITION 1 0 0`, your angle of view is parallel with the x-axis, so that the graph looks like a two-dimensional scatterplot of the z- and y-variables. At `VPOSITION 1 1 1`, you are looking down on the data box from a position just off the corner where the x-, y-, and z-values are highest.

The default is 1.5 −3.5 1.0.

Note The VPOSITION and VUP on page 1165 vectors must not be the same and must not differ by a constant multiplier.

VUP: Session subcommand for specifying which direction is up in relation to the data box

VUP K K K

Specifies which direction is up in relation to the data box (the view-up direction) as a ratio in x, y, and z. The default is `VUP 0 0 1`, which specifies that the positive z-direction is up. For example, `VUP 1 0 0` turns the data box on its side by making the positive x-direction up, instead of the positive z-direction.

Note The VPOSITION on page 1165 and VUP vectors must not be the same and must not differ by a constant multiplier.

WHISKER: Session subcommand for controlling the display of whiskers on a boxplot

WHISKER K

Controls the display of whiskers on a boxplot.

Value of K	Display of whiskers
0	Hides the whiskers.
1	Displays the whiskers.

WTITLE: Session subcommand for specifying the title of the output pane

WTITLE *"title"*

Specifies the title for the output pane.

You can use WTITLE as a subcommand with LAYOUT and all graph commands. The title you specify becomes the title of the output pane.

Model-Based Commands

FACPLOT: Session command for creating a factorial plot

FACPLOT C

FACPLOT "T|K"

Displays factorial plots for the model associated with the response variable in the specified column. To create this graph, you need to specify a model. To produce an interaction plot, you must have two or more variables.

Displays factorial plots for the model that you specify. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following code: `FACPLOT 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `FACPLOT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

To produce an interaction plot, you must have two or more variables. These plots can illustrate how a response variable relates to one or more variables. You can produce two types of factorial plots.

- Use a main effects plot to compare fitted means or data means across levels of one or more variables.
- Use an interactions plot to assess the effects of one variable at levels of another variable. An interactions plot displays the fitted means for a level of a variable with the level of a second variable held constant.

Although you can use these plots to display the effects, be sure to evaluate significance by looking at the effects in the analysis of variance table.

FACTORS C...C

Specifies the variables to display in both the main effect and interaction plots. You must specify 2 or more variables to generate an interaction plot.

GMAIN

GMAIN displays the main effect plot.

GINT

GINT displays the interactions plot.

FULL

Use FULL to display both interaction plots for each pair of categorical variables, such as Predictor 1 by Predictor 2 and Predictor 2 by Predictor 1. Sometimes you can identify interactions more easily when you examine both plots.

LOWER

Use LOWER to display one version of each interaction plot.

GMODEL (default)

GMODEL display plots for only terms that are in the model and listed in FACTORS. For purposes of a main effect plot, a variable is in the model if the model includes any terms that involve only that variable, e.g. A, A*A, or A*A*A. For purposes of an interaction plot, the interaction of two variables is in the model if the model includes any terms that involve both variables, e.g. A*B, A*B*B.

GALL

GALL displays plots for all terms involving variables on the FACTORS subcommand.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MFFCUBE: Session command for creating a cube plot for fitted means

MFFCUBE C

Creates a cube plot of the fitted means for the response in C.

You can draw a cube plot for a model from a 2-level factorial design. The factor columns can be numeric or text, and can contain any values. When the factor column is text, the first entry (row 1) is level 1, the next entry that is different is level 2. All factors must have 2 levels, and can have center points.

To draw a cube plot of the data means or a cube plot of the process variables in a mixtures design, use FFCUBE on page 341.

FACTORS C...C

Specifies the factors to include on the cube plot in C...C. You can specify 2–8 factors.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MMOPT: Session command for the Response Optimizer

MMOPT C...C

MMOPT "T|K"

You can use Minitab's Response Optimizer to identify the combination of input variable settings that jointly optimize a single response or a set of responses. All models in the optimization must be the same type of linear model (e.g. regression, binary logistic regression, general linear model, etc). Joint optimization must satisfy the requirements for all the responses in the set. The overall desirability (D) is a measure of how well you have satisfied the combined goals for all the responses. Overall desirability has a range of zero to one. One represents the ideal case; zero indicates that one or more responses are outside their acceptable limits.

You can use Minitab's Response Optimizer with a model in the worksheet to identify the combination of input variable settings that jointly optimize a single response or a set of responses. All models in the optimization must be the same type of linear model (e.g. regression, binary logistic regression, general linear model, etc). All models must be in the same worksheet. Joint optimization must satisfy the requirements for all the responses in the set. The overall desirability (D) is a measure of how well you have satisfied the combined goals for all the responses. Overall desirability has a range of zero to one. One represents the ideal case; zero indicates that one or more responses are outside their acceptable limits.

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `MMOPT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

Minitab calculates an optimal solution and draws a plot. The optimal solution serves as the starting point for the plot. This optimization plot allows you to interactively change the input variable settings to perform sensitivity analyses and possibly improve the initial solution.

Note Although numerical optimization along with graphical analysis can provide useful information, it is not a substitute for subject matter expertise. Be sure to use relevant background information, theoretical principles, and knowledge gained through observation or previous experimentation when applying these methods.

For each response, you must do the following:

- Select a goal.
- Specify a minimum acceptable value, a maximum acceptable value, and a target value.
- Specify a weight and importance level.

GOAL K...K

Specifies a goal for each response. Enter the goals in coded form, with 1 = "smaller is better", 2 = "target is best", 3 = "larger is better". "Smaller is better" means "make this response as small as possible". "Larger is better" means "make this response as large as possible". "Target is best" means "set this response as near to the target value as possible". You must enter one argument (K) for each model.

HOLD K1 K2 [K3]

Specifies a fixed level or range of levels for variable K1. The variables are indexed based on their order in the first model in which they appear. K2 without K3 specifies a fixed level. K2 and K3 together specify a range of levels.

HOLD levels should be specified in uncoded units.

Default for a factor is K2 = minimum value defined by models, K3 = maximum value defined by models.

Default for a covariate in a DOE model is a fixed level at the covariate data mean. Covariates can only be held at a constant level, not a range of levels.

UTWEIGHT K...K

Specifies the weights for each response. You must enter one argument (K) for each model. The weight defines the shape of the desirability function for each response. For each response, you can select a weight (from 0.1 to 10) to emphasize or de-emphasize the target. Weights place emphasis as follows:

- Weights that are less than 1 (the minimum is 0.1) place less emphasis on the target.
- Weights that are equal to 1 place equal importance on the target and the bounds.
- Weights that are greater than 1 (the maximum is 10) place more emphasis on the target.

The shape of the function as it goes from 0 to 1, or from 1 to 0, is determined by the weights value. If the value is 1, the shape is a straight line. If the value is less than 1, the shape is concave up. If the value is greater than 1, the shape is concave down.

IMPORTANCE K...K

Specifies the relative importance of each response in the composite desirability function. The number of arguments must match the number of models. These are numbers between 0.1 and 10. If all responses are equally desirable, use the default value of 1.0 for each response. The composite desirability is then the geometric mean of the individual desirabilities.

However, if some responses are more important than others, you can incorporate this information into the optimal solution by setting unequal importance values.

Larger values correspond to more important responses, smaller values to less important responses. You can also change the weights to perform sensitivity analyses on the optimal solution. For example, you may find that the optimal solution when one response has a large weight is very different from the optimal solution when the same response has a small weight.

START K...K

Specifies a starting point for the search algorithm by giving a coordinate for each factor. These are entered as constants. Specifying starting points is optional. However, if you specify any starting point, you must specify one value for each factor, and each value must be between the minimum and the maximum for that factor, or within the range of any user defined constraints.

DSTORE C

Stores the composite desirability values calculated at the points in DSTORE.

DLEVELS *E...E*

DLEVELS is required if DSTORE is used. This subcommand specifies the predictor levels in column or constants for the computation of the composite desirability stored in the DSTORE column. The number of arguments must match the number of predictors in all of the models. For categorical factors, the levels must be legal values of the corresponding factors.

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Minimum acceptable values, maximum acceptable values, and target values

Specifies the minimum acceptable values, the maximum acceptable values, and the target values for the responses. You must enter one argument (K) for each model. You must use all three subcommands because they are needed to obtain the numerically optimum solution. You must also enter a value for each response in each of these subcommands. In other words, you need a minimum acceptable value for each response, you need a maximum acceptable value for each response, you need a target value for each response.

If you have a "smaller is better" goal for a response, your goal is to minimize the response. Identify some point that can be described as the point of diminishing returns (i.e., while I want to minimize, going beyond a certain value makes little or no difference). If such a point exists, use it as the target value. If it does not, use a very small number, one that is probably not achievable, as the target value. For a "smaller is better" response, the target and the minimum acceptable value are also equal. The same can be said for a case where the goal is "larger is better", except that now you need a value on the upper end instead of the lower end of the range. When the goal is "target is best", you probably have upper and lower specification limits for the response, which can be used as minimum and maximum acceptable values. If not, you may want to use upper and lower points of diminishing returns, or some other method.

MINACCEPT K...K

Specifies the minimum acceptable values with K...K.

MAXACCEPT K...K

Specifies the maximum acceptable values with K...K.

TARGET K...K

Specifies the target values with K...K.

Output

OPTIPLOT

Displays the interactive optimization plot.

TPARAMETER

Displays the table of parameters that define the desirability to optimize.

TSOLUTION
Displays the table with the numerically optimal solution and the desirabilities for the responses.

THOLD
Displays the table with the hold values.

TSTART
Displays the table of user-specified start values for the search algorithm.

TPREDICTION
Displays the table of prediction with the confidence intervals.

NODEFAULT
Specifies that no default tables and graphs will be displayed.

MDIAGRAM: Session command to select multiple models for the response optimizer for commands from the Predictive Analytics module

MDIAGRAM "T|K" … "T|K"
Use a model diagram to combine multiple models from the Predictive Analytics commands into a single object. Then, you can use Minitab's Response Optimizer to identify the combination of input variable settings that jointly optimize the responses from those models.

T is the type of model. K is an index for the creation order of a model inside of the project. For example, `MDIAGRAM "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)
classification_treenet	TreeNet® Classification
classification_random_forest	Random Forests® Classification
regression_treenet	TreeNet® Regression
regression_random_forest	Random Forests® Regression
mars_regression	MARS® Regression

Minitab summarizes the models that you select and displays information about the predictors across the models.

Output

NODEFAULT

Specifies that no default tables and graphs will be displayed.

GIMPORTANCE K

Displays the variable importance chart.

The number of variables on the chart depends on the value of K. When K =

- 100: Displays all important variables (default).
- 0 < K ≤ 100: Displays a percentage of important variables. For example, `GIMPORTANCE 50` displays 50% of the important predictors.
- 0: Displays all predictor variables.

If multiple variables have the same importance and the percentage of variables to show includes at least 1, then the plot includes all the tied variables.

TMODEL

Displays the Model Performance table, which includes the response variable, a measure of model performance, and other information about the model. For a regression model, the measure of performance is an R-squared statistic. For a binary classification model, the measure of performance is the area under the ROC curve. For a multinomial classification model, the measure of performance is a misclassification rate. The measure of performance depends on the type of validation for the model.

TVARIABLE

Displays the Variable Ranges table, which includes the average importance, the range or categories, the predictor ID, and other information about the predictors.

GOPT: Session command for the Response Optimizer for commands from the Predictive Analytics module

GOPT "T|K"

You can use Minitab's Response Optimizer to identify the combination of input variable settings that jointly optimize a single response or a set of responses. Joint optimization must satisfy the requirements for all the responses in the set. The overall desirability (D) is a measure of how well you have satisfied the combined goals for all the responses. Overall desirability has a range of zero to one. One represents the ideal case; zero indicates that one or more responses are outside their acceptable limits.

To optimize multiple responses, first use the MDIAGRAM command to create a model diagram. Then, specify the model diagram as the type of model. To optimize a single response, specify the type and the number of the model in the project creation order.

Each type of model uses its own count. Model types are case-sensitive. For example, `GOPT "model_diagram|1";` produces results that jointly optimize all the responses for the responses in the first model diagram in the project. `GOPT "binary_logistic_regression|2";` produces results that optimize the response for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
model_diagram	Response Optimizer from the Predictive Analytics Module
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)
classification_treenet	TreeNet® Classification
classification_random_forest	Random Forests® Classification
regression_treenet	TreeNet® Regression
regression_random_forest	Random Forests® Regression
mars_regression	MARS® Regression

Minitab calculates an optimal solution and draws a plot. The optimal solution serves as the starting point for the plot. This optimization plot allows you to interactively change the input variable settings to perform sensitivity analyses and possibly improve the initial solution.

Note Although numerical optimization along with graphical analysis can provide useful information, it is not a substitute for subject matter expertise. Be sure to use relevant background information, theoretical principles, and knowledge gained through observation or previous experimentation when applying these methods.

For each response, you must do the following:

- Select a goal: minimize, maximize, or target.
- Specify a minimum acceptable value, a maximum acceptable value, and a target value. If the goal is to minimize, the specify the same value for the mimum acceptable value and the target. If the goal is to maximize, then specify the same value for the maximum acceptable value and the target.

Goal, minimum acceptable values, maximum acceptable values, and target values

Use the following subcommands to specify the goal, minimum acceptable value, target, and maximum acceptable value. You need a minimum acceptable value for each response, you need a maximum acceptable value for each response, and you need a target value for each response. You must enter one argument (K) for each response variable. If you have multiple response variables, specify the values for the subcommands in the same order as the model types from the MDIAGRAM command.

If you have a "smaller is better" goal for a response, your goal is to minimize the response. Identify some point that can be described as the point of diminishing returns (i.e., while I want to minimize, going beyond a certain value makes little or no difference). If such a point exists, use it as the target value. If it does not, use a very small number, one that is probably not achievable, as the target value. Specify the same value for the target and the minimum acceptable value.

If you have a "larger is better" goal, your goal is to maxmize the response. Identify some point that can be described as the point of diminishing returns (i.e., while I want to maximize, going beyond a certain value makes little or no difference). If such a point exists, use it as the target value. If it does not, use a very large number, one that is probably not achievable, as the target value. Specify the same value for the target and the maximum acceptable value.

When the goal is "target is best", you probably have upper and lower specification limits for the response, which can be used as minimum and maximum acceptable values. If not, you may want to use upper and lower points of diminishing returns, or some other method.

GOAL K...K

Specifies a goal for each response. Enter the goals in coded form, with 1 = "smaller is better", 2 = "target is best", 3 = "larger is better". "Smaller is better" means "make this response as small as possible". "Larger is better" means "make this response as large as possible". "Target is best" means "set this response as near to the target value as possible".

MINACCEPT K...K

Specifies the minimum acceptable values with K...K.

MAXACCEPT K...K

Specifies the maximum acceptable values with K...K.

TARGET K...K

Specifies the target values with K...K.

Optimal classes

OPTCLASS "T|K" "class"

OPTCLASS "T|K" "class"..."class"

OPTCLASS "T|K" "Event"

OPTCLASS "T|K" "Non-event"

Specifies the optimal class or classes for the search. If you do not specify the optimal class for a binary response, then the default optimal class is the event from when the model was fit. If you do not specify the optimal class for a multinomial response, then the default optimal class is the first class in value order for the column.

To specify the optimal class for a categorical variable, give the following information:

- the type of model (T)
- the number of the model in the project creation order (K)
- the classes for the algorithm to optimize with respect to the goal for that response variable

Use a separate instance of OPTCLASS for each categorical response variable. The type of model and the classes are case-sensitive.

For a binary response variable, specify one class. For a binary response variable where the model was fit with event/trial data, specify "Event" or "Non-event." The arguments for event/trial data are case-sensitive.

For a multinomial response variable, specify one or more classes.

For example, suppose a model diagram includes the following response variables:

- The first binary logistic regression model in the project was fit with the data in response/frequency format. The classes are success and failure. This subcommand optimizes success: `OPTCLASS "binary_logistic_regression|1" "success";`
- The second binary logistic regression model in the project was fit with data in event/trial format. The classes are success and failure. In the model diagram, the success class is the event. This subcommand optimizes success: `OPTCLASS "binary_logistic_regression|2" "Event";`
- The first TreeNet® classification model in the project has a multinomial response variable. The classes are red, yellow, green, and blue. This subcommand optimizes green and blue: `OPTCLASS "treenet_classification|1" "green" "blue";`

The subcommands for the analysis include the following:

Options

HOLD K1 K2 [K3]...[K]

By default, the algorithm searches continuous predictors through their range in the data and all levels of categorical predictors. Use this subcommand to limit the search to ranges for continuous variables or specific categories for categorical variables. K1 is the ID of the variable to hold. Use a separate `HOLD` subcommand for each variable that you want to limit during the optimization search. Predictor variables are identified based on their order in the first model in which they appear. The ID of the predictor is in the ID column of the Variable Ranges table of the model diagram.

For a continuous variable, K2 without K3 specifies a fixed value. K2 and K3 together specify a range. To treat a variable as missing, specify the variable ID and the missing value symbol, for example `HOLD 2 *`. Missing is not a valid hold level for a binary logistic regression model or a multiple regression model. Specify levels in uncoded units.

For a categorical variable, each K is a level to include in the search. To treat a variable as missing, specify the variable ID and two quotation marks, for example `HOLD 2 ""`. Missing is not a valid hold level for a binary logistic regression model or a multiple regression model. Specify levels in uncoded units.

UTWEIGHT K...K

Specifies the weights for each response. You must enter one argument (K) for each response variable. For multiple response optimization, enter the values in the order of the models from the `MDIAGRAM` command. The weight defines the shape of the desirability function for each response variable. For each response variable, you select a weight (from 0.1 to 10) to emphasize or de-emphasize the target. Weights place emphasis as follows:

- Weights that are less than 1 (the minimum is 0.1) place less emphasis on the target.
- Weights that are equal to 1 place equal importance on the target and the bounds.
- Weights that are greater than 1 (the maximum is 10) place more emphasis on the target.

The shape of the function as it goes from 0 to 1, or from 1 to 0, is determined by the weights value. If the value is 1, the shape is a straight line. If the value is less than 1, the shape is concave up. If the value is greater than 1, the shape is concave down.

IMPORTANCE K...K

Specifies the relative importance of each response variable in the composite desirability function. For multiple response optimization, enter the values in the order of the models from the `MDIAGRAM` command. The number of arguments must match the number of models. These are numbers between 0.1 and 10. If all responses are equally desirable, use the default value of 1.0 for each response. The composite desirability is then the geometric mean of the individual desirabilities.

Larger values correspond to more important responses, smaller values to less important responses. You can also change the weights to perform sensitivity analyses on the optimal solution. For example, you may find that the optimal solution when one response has a large weight is very different from the optimal solution when the same response has a small weight.

RANDOMIZE K

Specifies the base number for the random elements of the search for the genetic algorithm. K sets the starting point for Minitab's random number generator.

K must be a positive integer. Use 12345 to get the same results as Minitab's defaults.

STOPRULE

Specifies rules for when to stop the search for an optimal solution with one or more of the following subsubcommands. By default, the analysis uses the following specifications:

```
STOPRULE;
TIMEMINUTES 5;
MINDESIRE 1.
```

Use at least one of SEARCHES or TIMEMINUTES. Specify MINDESIRE after SEARCHES or TIMEMINUTES.

SEARCHES K

Specifies the number of times to iterate after the intial population for the genetic algorithm. Specify a non-negative integer. A value of 0 provides an optimal solution from the initial population.

TIMEMINUTES K

Specifies a maximum duration for the search as a number of minutes. Specify a non-negative number. A value of 0 provides an optimal solution from the initial population.

MINDESIRE K

Specifies to stop the search for an optimal solution if a solution reaches the minimum composite desirability. Specify a positive number less than or equal to 1.

Output

TSOLUTION K

Displays the table with the optimal solutions and the desirabilities. K specifies the number of solutions to include in the table.

IMPSHOW K

Specifies how many predictors ti display in the Solutions table. The table is in the output pane and at the bottom of the interactive optimization plot. The predictors in the table are in order of importance.

K = -1

Show all predictors, including unimportant predictors.

K = 0

Show only important predictors. This value is the default if the analysis does not include IMPSHOW.

K > 0

Specify an integer greater than 0 to display the top K predictors in order by importance. If the value is greater than the number of predictors, then the table shows all the predictors.

TPARAMETER

Displays the table of parameters for the search, including whether the goal is to minimize, maximize or target the response.

TPREDICTION

Displays the prediction table for the most optimal solution from the search.

THOLD

Displays the table with the hold values.

GICECURVE

Displays the interactive optimization plot.

Note The interactive optimization plot includes the table of the solutions at the bottom of the plot. To specify the number of solutions and number of predictors in the table use the TSOLUTIONS subcommand and the IMPSHOW sub-subcommand.

NODEFAULT

Specifies that no default tables and graphs will be displayed.

Storage

SSOL

Stores information about optimal solutions in rows of the worksheet, including predictor values, desirability values, and predicted values. Use NUMSOL to specify how many solutions to store, one solution per row. Use WORKSHEET to specify the name of the worksheet to use for storage.

NUMSOL K

Specify a positive integer to specify the number of solutions to store.

WORKSHEET *"text"*

WORKSHEET K

Specify the name of a worksheet to store the solutions in that worksheet. If a stored constant contains the name, then you can specify K.

Continuous response: `SDATA C C`

Categorical response: `SDATA C C C...C`

Multiple responses: `SDATA C C...C ... C...C`

Stores the predicted values and desirability values for rows of predictor values in the worksheet.

For a single response, the first column C is the column that stores the desirability from the model. The next column C specifies the column that stores the prediction. For a categorical response, specify an additional column for each level of the response variable. The columns store the probabilities for each level.

For multiple responses, specify a column for the composite desirability, then specify a set of columns for each response variable. The set includes the individual desirability for that response and the prediction. For a categorical response, the set also includes a column for each level of the response. The order of the sets is the order of the models from the MDIAGRAM command.

You can specify a range of columns. For example, suppose that you have data in columns c1-c10 and want to store the desirabilities and predictions in the next empty columns. The models are one continuous response and one binary response. Specify `SDATA c11-c17;` for a total of 7 columns. 1 column is for the composite desirability, 2 columns are for the continuous response, and 4 columns are for the binary response.

PREDICTORS C...C

Specifies the columns that have the predictor values for the calculation of the desirabilities. The order of the columns depends on the type of model. For a type of regression_multi_model or classification_multi_model from a Discover Best Model analysis, the order follows the method for the winning model. The following sections describe the orders for the model types.

Multiple responses from a model diagram

Specify the columns in ID order. For predictor variables from a model diagram, the variables are identified based on their order in the first model in which they appear. The ID of the predictor is in the ID column in the Variable Ranges table of the model diagram.

Linear regression or binary logistic regression

Specify the columns in the order of specification from the command, beginning with the continuous variables. For example, the subcommands that build the model have the following syntax:

```
CONTINUOUS C4 C5 C6;
CATEGORICAL C1 C3 C2;
```

Then specify the predictors in the following order:

```
PREDICTORS C4 C5 C6 C1 C3 C2;
```

TreeNet® classification, Random Forests® classification, TreeNet® regression, or Random Forests® regression

Specify the continuous predictors first in worksheet order. Specify the categorical predictors second, starting with the predictors with the fewest levels. For predictors with the same number of levels, follow the worksheet order. For example, the following table shows the worksheet order of the predictors and the number of levels for the categorical variables.

C1-T	C2	C3-T	C4	C5-T
Children, 2 Levels	Income, Continuous	Store, 3 Levels	Price, Continuous	ViewAd, 2 Levels
Yes	$37,000	A	$12.99	Yes
No	$47,000	B	$12.99	No
Yes	$34,000	C	$17.99	No

Specify these predictors in the following order:

```
PREDICTORS Income Price Children ViewAd Store;
```

MARS® Regression

Specify the categorical predictors in worksheet order from left to right. Then, specify the continuous predictors in worksheet order. For example, the following table shows the worksheet order of the predictors and the number of levels for the categorical variables.

C1-T	C2	C3-T	C4	C5-T
Children, 2 Levels	Income, Continuous	Store, 3 Levels	Price, Continuous	ViewAd, 2 Levels
Yes	$37,000	A	$12.99	Yes
No	$47,000	B	$12.99	No
Yes	$34,000	C	$17.99	No

Specfy these predictors in the following order:

```
PREDICTORS Children Store ViewAd Income Price;
```

MOVERCONT: Session command for creating an overlaid contour plot

MOVERCONT C...C

Use MOVERCONT to draw a contour plot for multiple responses. All models used in the overlaid contour plot must be the same type of linear model (for example, regression, binary logistic regression, general linear model, etc).

Contour plots show how a response variable relates to two variables based on a model equation. In a contour plot, the response surface is viewed as a two-dimensional plane where all points that have the same response are connected to produce contour lines of constant responses. Contour plots are useful for establishing operating conditions that produce desirable response values.

Applications that involve multiple responses present a different challenge than single response experiments. Since each response is important in determining the quality of the product, you need to consider the responses simultaneously. Optimal factor settings for one response may be far from optimal for another response. Overlaid contour plots allow you to visually identify an area of compromise among the various responses.

To produce this graph, you need to specify a model that contains at least 2 continuous variables. By default, the graph uses the first two continuous variables specified in the model as the x- and y-variables.

MOVERCONT uses the equation in the specified model to generate the z-values (responses) that appear in the graphs. The values used in the equation are the values in the x- and y-variables along with the HOLD and CHOLD values of the remaining variables.

You have the following contour plot options:

- Generating a single graph and choosing the x- and y-axis variables
- Specifying the values of contour levels on the graph

INUNIT K

Specifies the type of units to display the variables in the plots. Applies only to cases, such as DOE models, in which there are coded and uncoded units. Otherwise, the subcommand is ignored.

Value of K	Form of design
0	Coded
1	Uncoded

HOLD K...K

HOLD keeps the continuous variables at the settings you specify. HOLD is required if the model includes more than two continuous variables. With HOLD, you need to specify a constant for each continuous variables in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values are ignored for the variables that appear in each plot.

CHOLD K...K

CHOLD keeps the categorical variables at the settings that you specify. CHOLD is required if the model includes any categorical variables. With CHOLD, you need to specify one of the categorical levels for each categorical variable in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values are ignored for the variables that appear in each plot.

SINGLE K K

Specify a pair of variables to plot. K and K must be indices of continuous variables in the models. 1 is the first continuous variable, even when the categorical variables are specified first.

Each of these variables must appear in a term of at least one model. They do not need to appear in the same model.

The default plots the first two continuous variables that appear in model terms.

LEVELS K K...K K

Specifies the low and high value of the contour levels for each response in K K...K K. You must enter a pair of values for each response.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

MSURFACE: Session command for creating a surface plot

MSURFACE C

MSURFACE "T|K"

To produce this graph, you need to specify a model that contains at least 2 continuous variables. A surface plot shows how a response variable (the z-variable) relates to two variables (the x- and y-variables) based on a model. By default, the graph uses the first two continuous variables specified in the model as the x- and y-variables.

To produce this graph, specify a model that contains at least 2 continuous variables. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following command: `MSURFACE 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `MSURFACE "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

MSURFACE uses the specified model to generate the z-values (responses) that appear in the graphs. The values used are the values in the x- and y-variables along with the HOLD and CHOLD values of the remaining variables.

You have the following contour plot options:

- Generating a single graph and choosing the x- and y-axis variables
- Generating separate graphs for every combination of variables in the model

For information on how to change a surface plot to more easily compare it to contour plot, go to A comparison of MSURFACE and RMCONTOUR plots on page 1231.

For graphics options for MSURFACE, go to Graphics options for MSURFACE on page 1235.

INUNIT K

Specifies the type of units to display the variables in the plots. Applies only to cases, such as DOE models, in which there are coded and uncoded units. Otherwise, the subcommand is ignored.

Value of K	Form of design
0	Coded
1	Uncoded

HOLD K...K

HOLD keeps the continuous variables at the settings you specify. HOLD is required if the model includes more than two continuous variables. With HOLD, you need to specify a constant for each continuous variables in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values for the variables in the plot are ignored when you specify the default plot (of the first two variables) or SINGLE. When you specify ALLPAIRS, you generate many graphs. In each ALLPAIRS graph, the hold values are ignored for the variables that appear in the plot.

CHOLD K...K

CHOLD keeps the categorical variables at the settings that you specify. CHOLD is required if the model includes any categorical variables. With CHOLD, you need to specify one of the categorical levels for each categorical variable in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values for the variables in the plot are ignored when you specify the default plot (of the first two variables) or SINGLE. When you specify ALLPAIRS, you generate many graphs. In each ALLPAIRS graph, the hold values are ignored for the variables that appear in the plot.

ALLPAIRS

By default, MSURFACE uses the first two continuous variables specified in the model as x- and y-variables and produces one graph.

When you use ALLPAIRS, MSURFACE displays a separate graph for each possible combination of (x,y) factors with the calculated response (z). With n variables, using ALLPAIRS results in (n*(n - 1))/2 different contour

plots. When a graph of, for example, Temp as the x-variable and Pressure as the y-variable is generated, the transposed graph (that is, Pressure as the x-variable and Temp as the y-variable) is not generated.

You cannot specify both ALLPAIRS and SINGLE.

PANEL
PANEL displays all plots in panels of one graph.

SEPARATE
SEPARATE displays all plots in separate graphs.

SINGLE K K
When you specify SINGLE, MSURFACE displays a single graph. The two arguments are the continuous variables for the x- and y-axis, respectively.

LEVELS K K...K K
Specifies the low and high value of the contour levels for each response in K K...K K. You must enter a pair of values for each response.

The graphs are generated by calculating responses (z-values) using the values in the x- and y-variables and giving the other variables the HOLD or CHOLD values.

You cannot specify both ALLPAIRS and SINGLE.

TITLE *"title"*
Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

PREDICT: Session command for predicting response values

PREDICT C

PREDICT "T|K"
Predicts the response value of new observations for the model.

If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following command: `PREDICT 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `PREDICT "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)

T	Corresponding Analysis
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

Use a stored model to do the following:

- Predict the value of the response variable (fitted values) for the combinations of variable settings that you request.
- Create a confidence interval for the mean response at these combinations.
- Create a prediction interval for a single observation at these combinations (except for GZLM models).
- Store the prediction information in the worksheet.

CPREDICTORS C...C

Specifies new variable values with columns.

The number and type of arguments must match the variables in the model. In DOE, new values are assumed to be in uncoded units. For categorical variables, new values must be one of the values for the variable in the model.

Either CPREDICTORS or KPREDICTORS is required for all models. CPREDICTORS and KPREDICTORS are mutually exclusive.

KPREDICTORS K...K

Specifies new variable values with constants. KPREDICTORS can be repeated to add values for multiple new observations.

The number and type of arguments must match the variables in the model. In DOE, new values are assumed to be in uncoded units. For categorical variables, new values must be one of the values for the variable in the model.

Either CPREDICTORS or KPREDICTORS is required for all models. CPREDICTORS and KPREDICTORS are mutually exclusive.

CWEIGHT C (optional)

Use CWEIGHT when you use CPREDICTORS. CWEIGHT specifies new weight values with a column, one row for each combination of variables. Mixed effects models assume weights of 1 in prediction, even though the estimation of the model parameters can use weights that are different from 1.

The arguments must match the number of predictions to make.

CWEIGHT and KWEIGHT are mutually exclusive.

KWEIGHT K (optional)

Use KWEIGHT when you use KPREDICTORS. KWEIGHT specifies new weight values with constants. KWEIGHT can be repeated to add values for multiple new observations. The number of KWEIGHT commands must match the number of KPREDICTORS commands. Mixed effects models assume weights of 1 in prediction, even though the estimation of the model parameters can use weights that are different from 1.

The arguments must match the number of predictions to make.

CWEIGHT and KWEIGHT are mutually exclusive.

CCOVARIATES C...C

Either CCOVARIATE or KCOVARIATE can be used for models that contain covariates. If the model contains covariates but neither of these subcommands are issued, Minitab uses the mean of each covariate. CCOVARIATE and KCOVARIATE are ignored if the model is not a factorial design, general linear model, or a mixed effects model. For both subcommands, the number of arguments must match the number of covariates in the model.

Use CCOVARIATE when you use CPREDICTORS. CCOVARIATE specifies new covariate values with columns. The columns must be numeric.

COVARIATE and KCOVARIATE are mutually exclusive.

KCOVARIATES K...K

Either CCOVARIATE or KCOVARIATE can be used for factorial design models and general linear models that contain covariates. If the model contains covariates but neither of these subcommands are issued, Minitab uses the mean of each covariate. CCOVARIATE and KCOVARIATE are ignored if the model is not a factorial design, general linear model, or a mixed effects model. For both subcommands, the number of arguments must match the number of covariates in the model.

Use KCOVARIATES when you use KPREDICTORS. KCOVARIATE specifies new covariate values with constants. The constants must be numeric. KCOVARIATE can be repeated to add values for multiple new observations. The number of constants after KCOVARIATE must match the number of covariates in the model.

CCOVARIATE and KCOVARIATE are mutually exclusive.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

Results

Displays a variety of tables in the output.

NODEFAULT

Specifies that no default tables or graphs are displayed.

TEQUATION

Displays the model equation.

SINGLE

If you have categorical variables or factors, you can use SINGLE to display one equation instead of separate equations for each combination of categorical levels.

SEPARATE

Use SEPARATE to display separate equations for each level of each categorical variable. If you have only categorical variables, the display is always a single equation.

TPREDICTION

Displays the prediction table. In 2 cases, TPREDICTION displays the prediction table for any random level combinations defined by the random factors that are in the model:

- Stability study models with batch as a random factor
- Mixed effects models

TCPREDICTION

Displays the prediction table for selected levels of random factors in the model. TCPREDICTION is for 2 types of models:

- Stability study models with batch as a random factor
- Mixed effects models

TMODEL

Displays the model information table for general linear models and mixed effects models.

Storage

Stores information related to the predicted responses. The availability and function of the storage commands are different for mixed effects models and for stability study models that have batch as a random factor.

PFITS C

Stores the predicted responses in C. For stability study models with batch as a random factor, this subcommand stores the predicted responses for any random batch. For mixed effects models, this subcommand stores the predicted responses for any random level combinations defined by the random factors that are in the model.

CPFITS C

Stores the predicted responses for selected levels of a random factor in 2 cases:

- Stability study models with batch as a random factor
- Mixed effects models

PSEFITS C

Stores the standard error of the predicted responses in C. For stability study models with batch as a random factor, this subcommand stores the standard error of the predicted responses for any random batch. For mixed effects models, this subcommand stores the standard error of the predicted responses for any random level combinations defined by the random factors that are in the model.

CPSEFITS C

Stores the standard errors of the predicted responses for selected levels of a random factor in 2 cases:

- Stability study models with batch as a random factor
- Mixed effects models

DFCI C

For a mixed effects model, this subcommand stores the degrees of freedom for the confidence interval for the predicted responses at any random level combinations defined by the random factors that are in the model.

DFPI C

For a mixed effects model, this subcommand stores the degrees of freedom for the prediction interval for the predicted responses at any random level combinations defined by the random factors that are in the model.

CDFCI C

For a mixed effects model, this subcommand stores the degrees of freedom for the confidence interval for the predicted response at the selected levels of the random factors.

CDFPIC

For a mixed effects model, this subcommand stores the degrees of freedom for the prediction interval for the predicted response at the selected levels of the random factors.

CLIMITS C, C

Stores the lower and upper confidence limits of the predicted responses in C and C (except for GZLM models). For Stability Study models with batch as a random factor, this subcommand stores the confidence limits of the predicted responses for any random batch. For mixed effects models, this subcommand stores the confidence limits of the predicted responses for any random level combinations defined by the random factors that are in the model.

CCLIMITS C, C

Stores the confidence limits of the predicted responses for selected levels of a random factor in C and C in 2 cases:

- Stability study models with batch as a random factor
- Mixed effects models

PLIMITS C, C

Stores the lower and upper prediction limits of the predicted responses in C and C. For Stability Study models with batch as a random factor, this subcommand stores the prediction limits of the predicted responses for any random batch. For mixed effects models, this subcommand stores the prediction limits of the predicted responses for any random level combinations defined by the random factors that are in the model.

CPLIMITS C, C

Stores the prediction limits of the predicted responses for selected levels of a random factor in C and C in 2 cases:

- Stability study models with batch as a random factor
- Mixed effects models

BPFITS C

Stores the predicted responses using the original units for models that use the Box-Cox transformation. For Stability Study models with batch as a random factor, this subcommand stores the predicted responses in the original units for any random batch. Estimation for mixed effects models does not use the Box-Cox transformation, so the subcommand has no effect.

BCPFITS C

Stores the predicted responses for selected batches using the original units for models that use the Box-Cox transformation (Stability Study models with batch as a random factor only).

BCLIMITS C, C

Stores the confidence limits of the predicted responses using the original units for models that use the Box-Cox transformation. For Stability Study models with batch as a random factor, this subcommand stores the confidence

limits in the original units for any random batch. Estimation for mixed effects models does not use the Box-Cox transformation, so the subcommand has no effect.

BCCLIMITS C, C

Stores the confidence limits of the predicted responses for selected batches using the original units for models that use the Box-Cox transformation (Stability Study models with batch as a random factor only).

BPLIMITS C, C

Stores the prediction limits using the original units for models that use the Box-Cox transformation. For Stability Study models with batch as a random factor, this subcommand stores the prediction limits in the original units for any random batch. Estimation for mixed effects models does not use the Box-Cox transformation, so the subcommand has no effect.

BCPLIMITS C, C

Stores the prediction limits for selected batches using the original units for models that use the Box-Cox transformation (Stability Study models with batch as a random factor only).

MMPREDICT: Session command to calculate predictions from multiple models

MMPREDICT C...C

MMPREDICT "T|K"

The Response Optimizer uses MMPRED to create a table of with a prediction for the settings in the Response Optimizer.

The Response Optimizer uses MMPRED to create a table of predictions for multiple models with the same settings of the predictors. All the models must be created by the same command (for example, REGRESSION, GZLM, RSREG, etc). Models from GZLM must be the same type, all binary or all Poisson

For models in the same worksheet, you can predict for multiple models with the same settings of the predictors. All the models must be created by the same command (for example, REGRESSION, GZLM, RSREG, etc). Models from GZLM must be the same type, all binary or all Poisson.

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `MMPRED "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

Use a stored model to do the following:

- Predict the value of the response variable (fitted values) for the combinations of variable settings that you request.
- Create a confidence interval for the mean response at these combinations (if applicable for the model type).
- Create a prediction interval for a single observation at these combinations (if applicable for the model type).

To store the predictions in the worksheet, run the PREDICT command for each model.

KPREDICTORS K...K

Specifies new variable values with constants. The variables are indexed based on their order in the first model in which they appear.

The number and type of arguments must match the variables in the model. In DOE, new values are assumed to be in uncoded units. For categorical variables, new values must be one of the values for the variable in the model.

KCOVARIATES K...K

KCOVARIATE can be used for factorial design models and general linear models that contain covariates. If the model contains covariates but neither of these subcommands are issued, Minitab uses the mean of each covariate. KCOVARIATE is ignored if the model is not a factorial design, general linear model, or a mixed effects model. The number of arguments must match the number of covariates in the model.

KCOVARIATE specifies new covariate values with constants. The constants must be numeric. The number of constants after KCOVARIATE must match the number of covariates in the model.

Options

CONFIDENCE K

Specifies a confidence level. For example, for a 90% confidence level, enter `CONFIDENCE 90`. The default value of K is 95.

ITYPE K

Specifies the type of confidence interval.

Value of K	Type of confidence interval
–1	Lower bound
0 (default)	Two-sided
1	Upper bound

TITLE "title"

Specifies a title for the table. If you do not specify a title, Minitab uses a default title.

RMCONTOUR: Session command for creating a contour plot

RMCONTOUR C

RMCONTOUR "T|K"

To produce this graph, you need to specify a model that contains at least 2 continuous variables. A contour plot shows how a response variable (the z-variable) relates to two variables (the x- and y-variables) based on a model. By default, the graph uses the first two continuous variables specified in the model as the x- and y-variables.

To produce this graph, specify a model that contains at least 2 continuous variables. If the model is in a worksheet column, specify the column. For example, if a DOE model has a model in a column with the name *Strength* in the worksheet, specify the following command: `RMCONTOUR 'Strength';`

If the model is in the project, specify the type and the number of the model in the project creation order. Each type of model uses its own count. Model types are case-sensitive.

For example, `RMCONTOUR "binary_logistic_regression|2";` produces results for the second binary logistic regression model in the project.

The following table provides the valid model types.

T	Corresponding Analysis
multiple_regression	Linear Regression
binary_logistic_regression	Binary Logistic Regression
regression_multi_model	Discover Best Model (Continuous Response)
classification_multi_model	Discover Best Model (Binary Response)

The analysis is valid from a Discover Best Model analysis when the Linear Regression model or the Binary Logistic Regression model is the best model. If you specify a number for a model from Discover Best Model that is a different type, then the command produces an error.

RMCONTOUR uses the specified model to generate the z-values (responses) that appear in the graphs. The values used are the values in the x- and y-variables along with the HOLD and CHOLD values of the remaining variables.

You have the following contour plot options:

- Generating a single graph and choosing the x- and y-axis variables
- Generating separate graphs for every combination of variables in the model
- Specifying the number of contour levels on the graph
- Specifying the values of contour levels on the graph

For information on how to change a surface plot to more easily compare it to contour plot, go to A comparison of MSURFACE and RMCONTOUR plots on page 1231.

INUNIT K

Specifies the type of units to display the variables in the plots. Applies only to cases, such as DOE models, in which there are coded and uncoded units. Otherwise, the subcommand is ignored.

Value of K	Form of design
0	Coded
1	Uncoded

HOLD K...K

HOLD keeps the continuous variables at the settings you specify. HOLD is required if the model includes more than two continuous variables. With HOLD, you need to specify a constant for each continuous variables in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values are ignored for the variables that appear in each plot.

CHOLD K...K

CHOLD keeps the categorical variables at the settings that you specify. CHOLD is required if the model includes any categorical variables. With CHOLD, you need to specify one of the categorical levels for each categorical variable in the model.

For both HOLD and CHOLD, Minitab interprets the first value you enter as the hold value for the first variable that you entered in the model, the second as the second, and so on.

The hold values are ignored for the variables that appear in each plot.

ALLPAIRS

By default, MSURFACE uses the first two continuous variables specified in the model as x- and y-variables and produces one graph.

When you use ALLPAIRS, MSURFACE displays a separate graph for each possible combination of (x,y) factors with the calculated response (z). With n variables, using ALLPAIRS results in (n*(n – 1))/2 different contour plots. When a graph of, for example, Temp as the x-variable and Pressure as the y-variable is generated, the transposed graph (that is, Pressure as the x-variable and Temp as the y-variable) is not generated.

You cannot specify both ALLPAIRS and SINGLE.

PANEL

PANEL displays all plots in panels of one graph.

SEPARATE

SEPARATE displays all plots in separate graph.

SINGLE K K

When you specify SINGLE, MSURFACE displays a single graph. The two arguments are the continuous variables for the x- and y-axis, respectively.

NLEVELS K

Specifies number of contour levels. K must be a positive integer from 2 to 11.

LEVELS K...K

Specifies the values of the contour levels. The number of Ks must be from 2 to 11.

TITLE *"title"*

Specifies a title for the graph. If you do not specify a title, Minitab uses a default title.

Macros Session Commands

Structure Commands

GMACRO, MACRO, and ENDMACRO: Session commands for marking the beginning and ending of a macro

GMACRO

GMACRO must be the first line of your global macro. GMACRO specifies a global macro. GMACRO cannot be abbreviated.

MACRO

MACRO must be in the first line of your local macro, and specifies a local macro.

ENDMACRO

ENDMACRO ends all macros and must be in the last line of your macros. ENDMACRO cannot be abbreviated.

Declaration Statements

MCONSTANT, MCOLUMN, MMATRIX, and MTYPE: Session commands for declaring variables

All variables used in a local macro must be declared. Declaring a variable tells the local macro what type of variable to expect from the user, or the macro, while invoking.

Declaration requirements

- Declare variables that are constants with MCONSTANT, variables that are columns with MCOLUMN, and variables that are matrices with MMATRIX. (You may also use the plural synonyms MCONSTANTS, MCOLUMNS, and MMATRICES.) After the M- command, list all the variables that are of that type, separated by a space.
- An argument, which is a variable in the template, may be given the declaration MFREE. The variable data type—column, constant, or matrix—is determined by the type of the variable that is given when the macro is invoked. The macro statement MTYPE allows you to determine whether a variable declared with MFREE is a column, constant, or matrix.
- You may use a declaration statement several times, but only for different variables and only between the template and the body of the macro. Once a variable is declared, it cannot be redeclared. Variable declarations can only be made between the template and the body of the macro.
- The declaration commands (MCOLUMN, MCONSTANT, etc.) cannot be abbreviated.
- The declared variable must have a legal name. For more information, go to Using Variables.

Example of declaring variables

For example, suppose the template is as follows:

```
TRIM X Xbar
```

TRIM is the name of the macro and X and Xbar are variables that will be passed into the macro. The macro would need declaration statements that define whether X and Xbar are constants, columns, matrices, or "free" variables (defined below). Let's say X is a column in the global worksheet and Xbar is a constant in the global worksheet. The user would invoke the macro by typing, say, %TRIM C5 K1. The local macro file would have the following first few lines:

```
MACRO
TRIM X Xbar
MCOLUMN X
MCONSTANT Xbar
```

Note If you see an error that END does not follow READ or SET, a local variable could have the same name as a Minitab command. For example:

```
 SET col1
mini:maxi/1
END
```

where min and max are local variable names. Minitab interprets the second line as a command because MINI and MAXI are also Minitab commands. It displays the error message because it thinks you are trying to execute a command without first having entered the required END statement. You must avoid using Minitab commands for variable names if you need to use the variables in data entry.

Variable types

There are four special-purpose variables, which are each declared differently.

Subcommand
An implicit constant that has a value of either 1 (if the subcommand was invoked) or 0 (if the subcommand was not invoked).

Text
Declared with MCONSTANT as a text constant that contains a text string.

Suffixed
Declared with MCOLUMN or MCONSTANT as a range of columns or constants.

Free
Declared with MFREE as a column, constant, or matrix whose type is undetermined until the macro is invoked.

MFREE: Session command for declaring a free variable

You may want a local macro to operate with a column, constant, or matrix—whatever the user decides to use when he or she invokes the macro. The local macro can then take appropriate action, depending on the type of variable used when invoking the macro. A free variable is an argument variable whose type—column, constant, or matrix—is not determined until the macro is invoked.

Use a free variable in a macro

You must do five things in the local macro code to make free variables work:

1. List the free variable as an argument on the template. For example, here is a template for the macro TELLME that has X as an argument: TELLME X
2. Declare the free variable with the declaration statement MFREE. For example: MFREE X
3. Declare an additional variable as a constant: MCONSTANT Vartype

4. Use the macro statement MTYPE to analyze the free variable and store its variable type number in the constant declared in step. If the variable is a constant, then Vartype is set to 1; if it is a column, Vartype is set to 2; and if it is a matrix, Vartype is set to 3. You can include an MTYPE statement anywhere within the body of a local macro. For example, the command MTYPE X Vartype looks at the free variable X and stores its variable type (1, 2, or 3) in the constant Vartype.
5. Write code that can respond to the variable type that was used. In the following example, the IF statements make the macro perform different actions depending on what type of variable X is: IF Vartype = 1, NOTE X is a constant!, ELSEIF Vartype = 2, NOTE X is a column!, ELSE, NOTE X is a matrix!, ENDIF.

Invoke macros that use free variables like any other macros.

Note There is one case when the macro processor cannot determine the type of a variable. This happens when a variable that appears on an optional subcommand is declared as MFREE, and a user invokes the macro without using the subcommand. In this case, the macro processor assumes the variable is a column.

Example of a simple macro that uses free variables

The following local macro, TELLME, tells the user what kind of variable was used when the variable was invoked.

```
MACRO
TELLME X
MFREE X
MCONSTANT Vartype
MTYPE X Vartype
IF Vartype = 1
  NOTE X is a constant!
ELSEIF Vartype = 2
  NOTE X is a column!
ELSE
  NOTE X is a matrix!
ENDIF
ENDMACRO
```

You can invoke TELLME can be invoked in the following ways, which produce the following output:

Invoked like this	Produces this
%TELLME C1	X is a column!
%TELLME K1	X is a constant!
%TELLME M1	X is a matrix!

Example of a more complex macro that uses free variables

In the following local macro, BETWEEN.MAC, the arguments LOW and HI can be either columns or constants. Enter data in the worksheet before you run the macro so that you can specify X.1-X.N.

```
MACRO
BETWEEN X.1-X.N LOW HI ANS;
  STRICT.
MCOLUMN X.1-X.N L H ANS
MFREE     LOW HI
#
# X.1-X.N is a list of columns. LOW and HI can each be either
# a column or a constant.
#
# BETWEEN checks to see if the values in one row of X.1-X.N are
# all greater than or equal to LOW and all less than or equal
# to HI. If they are, the corresponding row of ANS is set 1.
# If not then ANS is set to 0. If the STRICT subcommand is used
# then BETWEEN checks for < and > rather than <= and >=.
#
RMINIMUM X.1-X.N L
RMAXIMUN X.1-X.N H
```

```
# Case where subcommand is not used
IF STRICT = 0
  LET ANS = ( L >= LOW ) AND ( H <= HI )
# Case where subcommand is used
ELSE
  LET ANS = ( L > LOW ) AND ( H < HI )
ENDIF
ENDMACRO
```

You can invoke BETWEEN in any of the following ways:

```
%BETWEEN C1-C3 .25 .35 C10

%BETWEEN C1-C3 C4 .35 C10

%BETWEEN C1-C3 .25 C5 C10

%BETWEEN C1-C3 C4 C5 C10
```

You can write a macro where a suffixed list of variables is declared as MFREE. But recall that all variables in a suffixed list must be of one type. Thus, in any one invocation of this macro, all the variables in the list must be of the same type. If you need to know what type of variable was passed in, use MTYPE.

DEFAULT: Session command for assigning default values to subcommand arguments

The DEFAULT statement is an optional line that allows you to assign a default value to a stored constant that appears on an optional subcommand. If a subcommand is not used when a user invokes the macro, the value on the DEFAULT line is used for the subcommand argument.

You cannot use DEFAULT to assign values to arguments on the main command - only arguments that are stored constants for a subcommand. Defaults for columns and matrices must be handled within the body of the macro.

Two rules about the syntax of DEFAULT:

- The DEFAULT line must come immediately after the declaration statements, before any other commands in the macro.
- The DEFAULT command cannot be abbreviated.

Local Macro Variables

KKCAT, KKNAME, and KKSET: Session commands for using text

You can use text data in columns, in stored constants, and as text strings in all three types of macros. In addition, you can pass a text string into a macro by enclosing the string in double quotes when invoking the macro. The passed string can then be assigned to a constant in your macro. Constants that hold text data are useful for specifying graph titles, file names, and names for variables that could be created in a local macro.

The following macro commands allow you to store text in a constant. They are especially useful for displaying titles and other annotation on macro output. The following text commands are used only in the body of global and local macros.

KKCAT K K K

Concatenates, or combines, the text in the first constant K with the text in the second constant K, and stores the combined string of text in the third constant K. For example, if the constant X contained "Mr." and the text constant Y contained "Jones", the following command KKCAT X Y Z would put the string "Mr.Jones" in constant Z.

KKNAME K C

Stores the name of column C in the constant K. For example, `KKNAME K1 C1` stores the name of column C1 in the constant K1.

KKSET K *"text"*

Stores the text within the double quotes in the constant K. You can also use the regular Minitab command LET to store text in constants. However, KKSET can store several text strings in several constants at once, whereas LET stores one text string in one constant. For example, `KKSET K1 "Text1" K2 "Text2"` stores the text strings in the constant K1 and K2.

Note In older versions of Minitab, you used single quotes around the text in KKSET. You can still use single quotes, but they are not recommended.

Example of a macro that uses text strings

The following local macro receives two strings when invoked and assigns them to constants.

```
MACRO
REVERSE file1 file2
#
# REVERSE reads the first 3 columns of the input file, file1.
#
MCONSTANT file1 file2
MCOLUMN X Y Z
PRINT file1 file2
#
# The FORMAT statement says that the data in file1 are text with 1 character.
# Valid rows would look like the following examples:
# a b c
# d e f
#
READ X Y Z;
  FORMAT(a1, x, a1, x, a1);
  FILE file1.
WRITE Z Y X;
  FILE file2.
#
# REVERSE now stores the 3 columns from file1 in reverse order as the output file,
file2.
#
ENDMACRO
```

Example of invoking a macro that uses text strings

We could use the preceding macro to reverse the columns in the file called INPUT.DAT and store the reversed data in the file called OUTPUT.DAT by using the following commands.

```
%REVERSE "INPUT" "OUTPUT"
```

DTYPE: Session command for determining the data type of a column or a constant

DTYPE *E* K

Use DTYPE to determine the data type of a column or constant (E), and store the results in a constant (K).

Value returned by DTYPE	Data type
0	Text
1	Real numbers
2	Integers
3	Date/time
10	Empty

DTYPE is often used with free variables (and the MFREE and MTYPE commands) in cases where the macro must be flexible enough to respond to a variety of possible inputs.

DTYPE is very useful when parts of your macro only work on some types of data. For example, you may have a subcommand of your local macro that lets the user specify a title for a graph; DTYPE can tell you if the user specified a text string or a number. Or, perhaps a part of your macro requires an integer; DTYPE could tell you if a variable was not an integer, allowing your macro to convert the real number to an integer.

Note DTYPE works only as a command. It does not work with IF or LET, for example.

Example of a macro that uses DTYPE

TELLDATA tells a user the data type of the variable that is specified when the macro is invoked.

```
MACRO
TELLDATA X
MFREE X
MCONSTANT Vartype
DTYPE X Vartype
IF Vartype = 0
  NOTE Variable is text
ELSEIF Vartype = 1
  NOTE Variable is real number
ELSEIF Vartype = 2
  NOTE Variable is integer
ELSEIF Vartype = 3
  NOTE Variable is date/time
#constants cannot be in date/time format
ELSEIF Vartype = 10
  NOTE Variable is empty
ENDIF
ENDMACRO
```

Control Statements

IF, ELSEIF, ELSE, ENDIF: Session commands for executing code depending on a logical condition

```
IF logical expression

ELSEIF logical expression

ELSE

ENDIF
```

Allows you to execute different blocks of code depending on a logical condition. A logical expression is any expression from the LET command. The comparison and Boolean operators listed below are the features of LET that are most often used in IF.

Comparison and Boolean operators	Alternative form	Description
=	EQ	Equal to
~=	NE	Not equal to
<	LT	Less than
>	GT	Greater than
<=	LE	Less than or equal to
>=	GE	Greater than or equal to
&	AND	And
|	OR	Or
~	NOT	Not

In most cases the logical expression evaluates to a single number. If the number is 0 (false), the block of statements is skipped; if it is not 0 (true), the block is executed. If the logical expression evaluates to a column, then if all entries in the column are 0, the expression is considered false, otherwise it is considered true.

You can use multiple ELSEIF statements within the IF-ENDIF block.

The following is a simple example, using a global macro. Enter data in columns C1-C3 before you run the macro.

```
GMACRO
SMALL
#
# Takes the data in C1-C3. Finds the column with the smallest mean
# and prints that column. If, because of ties, there is no single column
# with the smallest mean, a message is printed.
#
LET K1 = MEAN(C1)
LET K2 = MEAN(C2)
LET K3 = MEAN(C3)
IF K1 < K2 AND K1 < K3
  PRINT C1
ELSEIF K2 < K1 AND K2 < K3
  PRINT C2
ELSEIF K3 < K1 AND K3 < K2
  PRINT C3
ELSE
  NOTE Note: There are ties.
```

```
ENDIF
ENDMACRO
```

DO and ENDDO: Session commands for looping through a block of commands

DO K

ENDDO

Allows you to loop through a block of commands. K is set equal to the first number in the list, then the block of commands is executed. When Minitab reaches the ENDDO, K is set equal to the next number in the list and the block is executed again. This continues until all numbers in the list are used, or until you branch out of the DO-loop with a BREAK, GOTO, RETURN, or EXIT command.

The list of numbers can be an explicit list of any numbers or stored constants. A patterned list can be abbreviated using a colon and slash as in SET. For example, 1:10 is the list 1, 2, 3, ... , 10, and 1:1.8 /.2 is the list 1, 1.2, 1.4, 1.6, 1.8. Numbers can be in either increasing order or decreasing order. The following DO-loop changes the values in rows 1 through 10 and row 50 of columns C1 and C2 to the missing value code:

```
DO  K1 = 1:10 50
  LET C1(K1) = '*'
  LET C2(K1) = '*'
ENDDO
```

The following is a local macro that calculates a moving average of length three. It shows how to loop through the values in a column. Enter data in a column of the worksheet before you run the macro so that you can specify X.

```
MACRO
MOVAVE  X  Y
#
# Calculates the simple moving average of the data in X and
# stores the answer in Y.
#
MCONSTANT N  I
MCOLUMN   X  Y
LET N = COUNT(X)
LET Y(1) = '*'
LET Y(2) = '*'
DO I = 3 : N
  LET Y(I) = (X(I) + X(I-1) + X(I-2))/3
ENDDO
ENDMACRO
```

Note Instead of modifying a worksheet variable inside a DO/ENDDO loop, copying the worksheet variable to a local macro variable, modifying the macro variable in the loop, then copying the macro variable back to the worksheet variable might be faster.

WHILE and ENDWHILE: Session commands for repeating a block of commands depending on a logical expression

WHILE *logical expression*

Repeats a block of commands as long as a logical expression is true.

ENDWHILE

Marks the end of the WHILE loop.

Repeats a block of commands as long as the logical expression is true. The logical expression follows the same rules as in the IF statement.

Suppose you want to find the root of the equation, $y = -1 + x + x^3$. This equation has only one real root, which is between 0 and 1. The following global macro calculates, approximately, what the root is.

```
GMACRO
ROOT
#
# Finds the root of a specific polynomial. The result is
# within .01 of the exact answer.
# K90-K93 are used for scratch work
#
NAME  K90 'X'  K91 'Y'  K92 'Xlow'  K93 'Ylow'
LET 'X' = 0
LET 'Y' = -1
WHILE 'Y' < 0
  LET 'X' = 'X' + .01
  LET 'Y' = -1 + 'X' + 'X'**3
ENDWHILE
LET 'Xlow' = 'X' - .01
LET 'Ylow' = -1 + 'Xlow' + 'Xlow'**3
PRINT 'Xlow' 'Ylow'  'X' 'Y'
ENDMACRO
```

The macro first initializes the two variables, X and Y, to 0 and −1. Each time through the WHILE-loop, Minitab first checks to see that Y is still less than zero. If it is, Minitab increases X by .01 and calculates Y at this new value. When the condition fails—that is, when Y is no longer less than zero—the macro exits the loop and goes to the first statement after ENDWHILE. Then, the macro prints the result.

Note Instead of modifying a worksheet variable inside a WHILE / ENDWHILE loop, copying the worksheet variable to a local macro variable, modifying the macro variable in the loop, then copying the macro variable back to the worksheet variable might be faster.

NEXT: Session command for transferring control from a loop to the beginning of the block

Transfers control from within a DO- or WHILE-loop back to the beginning of the block. For DO, the loop variable is then set to the next value in the list and the loop is executed again. The following is a simple example, using a global macro.

```
GMACRO
FIVES
#
# Takes the column named X and changes all entries
# that are greater than 5 to 5.
# Constants K90 and K91 are used for scratch work.
#
NAME K90 'N' K91 'I'
LET 'N' = COUNT('X')
DO 'I' = 1 : 'N'
  IF  'X'('I') <= 5
    NEXT
  ELSE
    LET 'X'('I') = 5
  ENDIF
ENDDO
ENDMACRO
```

The DO-loop goes through all the values in X. If a value is less than or equal to 5, NEXT passes control to the top of the DO-loop and the value is left unchanged. If a value is greater than 5, the ELSEIF block is executed and that value is set to 5.

BREAK: Session command for transferring control from a DO- or WHILE-loop

Transfers control from within a DO- or WHILE-loop to the command immediately following the end of the loop. Thus BREAK breaks out of the loop.

The following is a simple example of BREAK in a global macro. The program works on a worksheet where one of the columns has the name X. The program goes through the values of X until it finds a missing value. It then leaves the loop and goes to the statement following ENDDO—in this example, DELETE. Note that this program does not handle the case when X has no missing values correctly. For an example that handles the case when X has no missing values, go to EXIT: Session command for transferring control back to Minitab or for closing Minitab on page 1203.

```
GMACRO
NOMISS
#
# Takes data from the column named X. Finds the first missing
# observation. Then deletes all observations starting with the
# first missing to the end of the column.
# Constants K90 and K91 are used for scratch work
#
LET K90 = COUNT('X')
DO K91 = 1:K90
  IF  'X'(K91) = '*'
    BREAK
  ENDIF
ENDDO
DELETE  K91:K90 'X'
ENDMACRO
```

GOTO and MLABEL: Session commands for branching to any line in a macro

GOTO *number*

MLABEL *number*

Allows you to branch to any line in your macro. There can be several GOTO's in one program. A GOTO is matched to the MLABEL that has the same number. The number can be any integer from 1 to 8 digits long. It cannot be a variable.

The following example is a modification of the macro NOMISS, but uses GOTO instead of BREAK. The program works on a worksheet where one of the columns has the name X. The program goes through the values of X until it finds a missing value. It then leaves the loop and goes to the statement following ENDDO—in this example, DELETE. If X has no missing values, the program prints a note.

```
GMACRO
NOMISS3
#
# Takes data from the column named X. Finds the first missing
# observation. Then deletes all observations starting with the
# first missing to the end of the column.
# Constants K90 and K91 are used for scratch work
#
LET k90 = COUNT('X')
DO K91 = 1:k90
  IF  'X'(K91) = '*'
    GOTO 5
  ENDIF
  IF k91 = k90
  NOTE Note: there are no missing observations in X
  Exit
  ENDIF
```

```
ENDDO
MLABEL 5
DELETE  K91:k90 'X'
ENDMACRO
```

CALL and RETURN: Session commands for passing control to another macro

CALL *template*

RETURN

You can include several macros in one file, just as a program often includes several subroutines. CALL and RETURN let you specify when to pass control to another macro and when to return to the main macro. You can include several global macros in one file, or several local macros in one file, but you cannot mix global and local macros together in one file.

When you invoke a macro, from interactive Minitab or from another macro, the first macro in the file is executed first. Use the macro statements CALL and RETURN to invoke a different macro within the macro file.

Recall that the second line of a macro is the template, or the macro name. When one macro in a macro file calls another macro in that file, use the command CALL, followed by the name on that macro's template. If it is a local macro, include appropriate arguments and subcommands. Any macro in a macro file can CALL any other macro in the file, any number of times.

RETURN says to leave the current macro and go back to the calling macro, to the statement just after the CALL. RETURN is optional. If RETURN is not present in the macro that was called (the subroutine), then, after it has executed, control is transferred back to the calling macro.

The following example is a variation on ANALYZE2.MAC. This example, ANALYZE3, uses the TSET command to ask the user whether to print all the data. If the response is "yes", then the macro sets K80 to 1. If the answer is anything else, then the macro sets K80 to 0. For more information, go to READ, TSET, and SET: Session command for asking users questions and using the answers in a macro on page 1211. The OK subroutine checks the value of K80 with an IF statement. If K80 equals 1, then the RETURN statement sends control back to the main macro. If K80 is anything else, then the macro prints one more note. When the ENDMACRO statement is encountered in either the TOOSMALL or OK subroutine, control is transferred back to the calling macro.

```
GMACRO
ANALYZE3
#
MTITLE "ANALYZE3"
NOTE Do you want all the data printed?
NOTE Type "yes" or "no" in quotation marks,
NOTE then click the Submit button.
tset c10;
file "terminal".
IF c10(1) = "yes"
 LET k80 = 1
ELSE
 LET k80 = 0
ENDIF
ERASE c10.
# If user types "yes" K80 = 1, if "no" K80 = 0
LET K90 = COUNT(C1)
IF K90 < 5
  CALL TOOSMALL
ELSE
  CALL OK
ENDIF
#
IF K80 = 1
NOTE Here are the data.
```

```
PRINT C1-C3
ENDIF
ENDMTITLE
ENDMACRO
#
#
GMACRO
TOOSMALL
NOTE Data set has fewer than 5 observations.
NOTE No analysis will be done.
ENDMACRO
#
#
GMACRO
OK
NAME C1 'Yield' C2 'Chem1'  C3 'Chem2' C5 'Ln.Yield'
DESCRIBE C1-C3
LET C5 = LOGE('Yield')
REGRESS;
RESP c5;
CONT c2 c3;
TERMS c2 c3.
IF K80 = 1
  RETURN
ENDIF
NOTE Analysis done, but no data printed by request
ENDMACRO
```

EXIT: Session command for transferring control back to Minitab or for closing Minitab

EXIT

EXIT has two very different behaviors depending on whether it is used in global and local macros, or in an exec file, as follows:

- In a global or local macro, EXIT transfers control back to interactive Minitab.
- In an exec file, EXIT closes Minitab.

The following example is a modification of the macro NOMISS, which correctly handles the case when X contains no missing values. The program works on a worksheet where one of the columns has the name X. The program goes through the values of X until it finds a missing value. It then leaves the loop and goes to the statement following ENDDO—in this example, DELETE. If X has no missing values, the program prints a note.

```
GMACRO
NOMISS2
#
# Takes data from the column named X. Finds the first missing
# observation. Then deletes all observations starting with the
# first missing to the end of the column. Prints a message if
# the column has no missing values.
# Constants K90 and K91 are used for scratch work
#
LET K90 = COUNT('X')
DO K91 = 1:K90
  IF  'X'(K91) = '*'
    BREAK
  ENDIF
  IF K91 = K90
    NOTE Note: There are no missing observations in X.
    EXIT
  ENDIF
ENDDO
DELETE  K91:K90 'X'
ENDMACRO
```

PAUSE and RESUME: Session commands for pausing and resuming a macro

PAUSE

When Minitab encounters a PAUSE in a macro, control is shifted from the macro to the keyboard. You can then type any Minitab command. PAUSE can help you debug a macro you are developing. It can also allow you to get input from the macro user.

If you are in PAUSE mode from within a local macro, you have access to the local worksheet and only the local worksheet. You can also declare new local variables and use them. They will be stored at the end of the local worksheet.

When you are in PAUSE mode, you can type any Minitab command. You cannot CALL other macros in the same file, invoke a macro from another macro file, or use control statements.

RESUME

When you want to return control to the macro, type RESUME (or just R).

Using DOS Commands

CD: Session command for displaying or changing the current directory

CD ***[filepath]***

CD without a path displays the current directory. CD with a path changes the current directory to the one that you specify.

For example, `CD` displays the current directory, and `CD WILLIAMS\SALES91` changes the current directory to WILLIAMS\SALES91.

Labeling Macro Output

MTITLE: Session command for adding a title above output

MTITLE ***"title"***

Starts the `MTITLE` mode. This mode adds a title for any output that is produced by commands between `MTITLE` and `ENDMTITLE`. While in `MTITLE` mode, you cannot save a project, open a project, open a worksheet, or open a graph. You cannot nest `MTITLE` in another `MTITLE` command. Because you cannot nest `MTITLE`, you usally do not use `MTITLE` in a macro that is nested in another macro.

NOTITLE

Suppresses the titles of the commands executed while in `MTITLE` mode.

ENDMTITLE
Ends the MTITLE mode. Minitab does not display any output until you specify ENDMTITLE.

TITLE: Session subcommand for adding a title to a graph

TITLE *["title"]*
Adds a title to a graph.

- You can use TITLE more than one time. Each additional title goes below the preceding title.
- Using TITLE with no argument displays the default title. Using TITLE with an argument replaces the default title.
- Default titles precede additional titles.

TITLE works the same for the basic graph commands and control chart commands. You can suppress the default title with the subcommand NODTITLE.

Attributes for the title text

TFONT *"font name"*
TFONT specifies the font name for the text.

TCOLOR K
TCOLOR specifies the text color, which is a number from Numbers for colors to use in session commands on page 1255.

TSIZE K
TSIZE specifies the size of the text in Minitab units. K can be any positive number.

PSIZE K
PSIZE specifies the size of the text in points. K can be any positive number.

BOLD
Specifies to display the text in bold font.

NOBOLD
Specifies to not display the text in bold font.

UNDERLINE
Specifies to underline the text.

NOUNDERLINE
Specifies to not underline the text.

ITALIC
Specifies to display the text in italic font.

NOITALIC
Specifies to not display the text in italic font.

Positioning

LEFT
Puts the title in the left of the figure region.

CENTER (default)
Puts the title in the center of the figure region.

RIGHT
Puts the title in the right of the figure region.

OFFSET K K
Moves a title away from the position specified with LEFT, CENTER or RIGHT. The first argument specifies how far to move the footnote horizontally (in figure units). The second argument specifies the how far to move the footnote vertically (in figure units). The default offset is 0.0 0.0.

WTITLE: Session subcommand for specifying the title of the output pane

WTITLE *"title"*
Specifies the title for the output pane.

You can use WTITLE as a subcommand with LAYOUT and all graph commands. The title you specify becomes the title of the output pane.

Debugging Tools

ECHO and NOECHO: Commands for displaying Minitab commands from macros and execs

The ECHO and NOECHO commands control whether commands in a macro or exec are in the output. When you develop a macro, you can use ECHO to see the commands so that you can find errors more easily.

You can submit ECHO and NOECHO before you invoke a macro. You can also place them anywhere within the body of a macro. You can use ECHO and NOECHO several times in a macro to turn on and off the display of commands.

ECHO
In ECHO mode, only commands in the body of the macro are in the output. The commands include Minitab commands, macro statements, and invocations of macros in other files. The commands do not include the template and declarations. (Declarations are in local macros only.) Text that is after a # is not in the output.

NOECHO (default)
In NOECHO mode, no Minitab commands or macro statements are in the output. The output of Minitab commands is in the output.

PAUSE and RESUME: Session commands for pausing and resuming a macro

PAUSE

When Minitab encounters a PAUSE in a macro, control is shifted from the macro to the keyboard. You can then type any Minitab command. PAUSE can help you debug a macro you are developing. It can also allow you to get input from the macro user.

If you are in PAUSE mode from within a local macro, you have access to the local worksheet and only the local worksheet. You can also declare new local variables and use them. They will be stored at the end of the local worksheet.

When you are in PAUSE mode, you can type any Minitab command. You cannot CALL other macros in the same file, invoke a macro from another macro file, or use control statements.

RESUME

When you want to return control to the macro, type RESUME (or just R).

Additional Local Macro Features

INFO: Session command for summarizing the current worksheet

INFO [C...C]

Summarizes the current worksheet.

If no columns are specified, INFO prints a list of all columns used with their names and counts, all stored constants, all matrices. If there are missing observations, a count of these is also given. If a column contains text data, the letter T is printed to the left of the column. If columns have assigned formulas, these are printed along with the method selected for updating the calculations (manual or automatic). If you list columns, information is given on just those columns.

WRITE: Session command for storing data in a text file

WRITE *E...E*

Writes data in the specified columns or constants to the screen or to a data file.

Because of potential conflicts with the global worksheet, the commands SAVE and RETRIEVE do not work in a local macro. Global worksheet variables that have been passed into the macro as arguments assume any new values given to them during the course of the macro execution. You can always save those variables after the macro executes. But you may also want to save local worksheet variables that are not passed as arguments. You can use WRITE to save local worksheet variables, use the WRITE command within your macro.

FILE *"filename"*

Specifies the file to store the data in.

Example

Suppose you have three column variables in the local worksheet named X, Y, and Z. The following command saves those three columns in a text file named MYWORK.DAT.

```
WRITE X Y Z;
  FILE "MYWORK".
```

Commands that Affect Output

BRIEF: Session command for controlling the amount of output

BRIEF K

Controls the amount of output. For example, the following table describes how BRIEF works with most commands that create a designed experiment.

Value of K	Output that is displayed
0	Minitab displays no output from the command, but performs all specified storage and displays the following output: error messages, warnings, prompts, and notes; graphs; WRITE to the screen.
1	Minitab displays a summary of the design.
2 (default)	Same as K = 1 output.
3	Same as K = 2 output, but Minitab also displays the design table.

Used as a main command, BRIEF affects the amount of output produced by subsequent commands. Used as a subcommand, BRIEF only affects output for the command it is used with.

Most commands are affected by BRIEF only when it is set to 0. However, BRIEF affects the amount of output produced by the following commands in specific ways. Go to the individual commands for details.

ARIMA on page 934

BBDESIGN on page 343

CCDESIGN on page 345

CLUOBS on page 915

CLUVARS on page 919

DISCRIMINANT on page 922

EVDESIGN on page 367

FACTOR on page 912

FFDESIGN on page 280

KMEANS on page 921

LREGRESSION on page 806

LTABLE on page 789

LTEST on page 771

MIXREG on page 377

NLOGISTIC on page 217

OLOGISTIC on page 214

OPTDES on page 363

PROBIT on page 820

SCDESIGN on page 371

SLDESIGN on page 374

GSAVE: Session subcommand for saving a graph in a file

GSAVE *"file_name"*

GSAVE K

Saves the graph in a file.

The default file name is Minitab.PNG. You can specify a custom file name in double quotation marks ("file_name"), or as a stored text constant (K). You can also use any of the following subcommands to save the graph in a different graphics format.

Some graph commands—for example, HISTOGRAM C1 C2 C3—generate more than one graph. If you include the GSAVE subcommand with such a command, Minitab saves multiple files. Minitab gives each file a different file name. Minitab uses the first five characters of the name you specify, then appends a number (001, 002, and so on), for up to 300 files.

JPEG
JPEG color

PNGB
PNG grayscale

PNGC
PNG color

TIFB
TIF grayscale

TIF
TIF color

BMPB
BMP grayscale

BMPC
BMP color

GIF
GIF

EMF
EMF

RESOLUTION K
Saves the graph at a resolution of K dots per inch.

GSCALE: Session command to determine appropriate scaling for a graph

GSCALE K K
GSCALE is useful primarily when you are writing a macro that produces graphs, and you need to know information before you produce the graphs to ensure that the scaling on the graphs will look right. For example, you might want to generate two or more graphs that use the same scale, but you want to control the scale. You can use the data stored by GSCALE to specify scaling options in subsequent graph commands.

The arguments on the main command are the minimum (the first K) and the maximum (the second K) of the data from the columns to be graphed, combined. An easy way to get those values is to STACK all of the columns on top of each other in a new column, then use the MIN and MAX commands to store the minimum and maximum values.

NMINIMUM K
Specifies the minimum number of ticks to use.

NMAXIMUM K
Specifies the maximum number of ticks to use.

NTICKS K
Stores the number of ticks.

TMINIMUM K
Stores the minimum tick value.

TMAXIMUM K
Stores the maximum tick value.

TINCREMENT K
Stores the distance between ticks.

SMINIMUM K
Stores the scale minimum.

SMAXIMUM K
Stores the scale maximum.

NOBRUSH: Session subcommand for disabling brushing on a graph

Can be used as a subcommand of any graphics command to disable brushing on the resulting graph. Why disable brushing? Brushing can only highlight rows of data in the global worksheet. But graphs created in local macros are sometimes based on data in the local worksheet that have no relationship to corresponding rows of data in the global worksheet.

Communicating with Macro Users

READ, TSET, and SET: Session command for asking users questions and using the answers in a macro

READ, TSET and SET have a special feature that allows you to ask users questions and then use their answers in the macro. A macro will pause for user input if you use READ, TSET or SET with the subcommand FILE with the special filename TERMINAL. TERMINAL tells Minitab to wait for input from the keyboard.

READ, TSET and SET have other subcommands.

```
READ C...C
    FILE "TERMINAL"

SET C
    FILE "TERMINAL"

TSET C
    FILE "TERMINAL"
```

PAUSE and RESUME: Session commands for pausing and resuming a macro

PAUSE

When Minitab encounters a PAUSE in a macro, control is shifted from the macro to the keyboard. You can then type any Minitab command. PAUSE can help you debug a macro you are developing. It can also allow you to get input from the macro user.

If you are in PAUSE mode from within a local macro, you have access to the local worksheet and only the local worksheet. You can also declare new local variables and use them. They will be stored at the end of the local worksheet.

When you are in PAUSE mode, you can type any Minitab command. You cannot CALL other macros in the same file, invoke a macro from another macro file, or use control statements.

RESUME

When you want to return control to the macro, type RESUME (or just R).

NOTE: Session command for adding comments that are displayed in the output

NOTE

Use NOTE to annotate your macro program with comments that are displayed in the output. To annotate with comments that are not displayed, use the comment symbol #.

Use notes to make your output more readable with spaces or to add helpful notes to yourself. Text from notes will appear in separate output tabs unless you use the MTITLE and ENDMTITLE commands to specify a block of output.

- Put the NOTE command at the beginning of a line.
- All text on that line will be ignored by the macro processor. However, text on a NOTE line (except the first five spaces – the word NOTE and a space) does display in the output when the macro is executed.

Note The NOTE cannot end with a semicolon. You can use a semicolon, but end the line with it. For example, NOTE AB;CD is acceptable, but NOTE ABCD; is not acceptable.

You can also make your macro file more readable by adding blank lines between the lines of macro statements and commands. The blank lines do not interfere with the execution of the macro, and will not appear in the output. You do not have to start a blank line with the comment symbol #.

Example of using NOTE in a macro

Macro Code	Results
GMACRO SAMPLEMACRO RAND 50 C1 C2 C3; UNIFORM 5 10. MTITLE "Print Columns". NOTE Here come the data NOTE PRINT C1-C3. ENDMTITLE ENDMACRO	Here come the data **Data** Row Yield Chem1 Chem2 1 11.28 87 1.83 2 8.44 61 25.42 3 13.19 59 28.64 ...
Note If you don't use MTITLEand ENDMTITLE, then this example will generate 2 tabs of output. The instance of NOTE without any text does not generate an output tab.	1. The first tab will contain Here come the data. 2. The second tab will be blank. 3. The third tab will contain the data.

Execs Commands

ECHO and NOECHO: Commands for displaying Minitab commands from macros and execs

The ECHO and NOECHO commands control whether commands in a macro or exec are in the output. When you develop a macro, you can use ECHO to see the commands so that you can find errors more easily.

You can submit ECHO and NOECHO before you invoke a macro. You can also place them anywhere within the body of a macro. You can use ECHO and NOECHO several times in a macro to turn on and off the display of commands.

ECHO

In ECHO mode, only commands in the body of the macro are in the output. The commands include Minitab commands, macro statements, and invocations of macros in other files. The commands do not include the template and declarations. (Declarations are in local macros only.) Text that is after a # is not in the output.

NOECHO (default)

In NOECHO mode, no Minitab commands or macro statements are in the output. The output of Minitab commands is in the output.

EXECUTE: Session command for running an Exec file

EXECUTE *"filename"* K

Note You cannot use EXECUTE in a local macro.

Runs commands that are stored in a file. These command files are called Execs.

You may specify the filename as either the name of the file in double quotes, or a stored text constant. The default file extension for Execs is MTB. When you use EXECUTE, you do not need to type the file extension if it is .MTB. The default file name is Minitab.MTB. If you do not specify a file name with EXECUTE, Minitab runs Minitab.MTB if it exists.

The optional argument K lets you specify how many times to run the Exec. K can be any integer. The default value is 1, which means that the Exec runs one time. If K > 1, the Exec runs K times. If K is < 1, the macro does not run. For information on using K to determine whether to run an Exec, go to Using Conditional Execution

To interrupt an Exec, press **Ctrl+Break**. Minitab will finish the command in process before it stops the macro.

Supporting Concepts

Add your own function

1. Start a word processor or text editor, such as Notepad.
2. Open the file USERFUNC.MAC. The file is in the Macros folder of your main Minitab folder.
3. Back up the file by saving it as a different name, such as USERFUNC.BAK. Save the file as plain text (ASCII file).
4. In USERFUNC.MAC, add your own function by following the instructions at the top of the file. In the macro you will assign your function a number between 1 and 1000.
5. Save USERFUNC.MAC as a plain text file.

Adding comments to a macro

You can add comments to a macro that are displayed or not displayed.

Add comments that are displayed

Use the NOTE on page 1211 command at the beginning of a line that contains the comment.

Add comments that are not displayed

Place the symbol # anywhere on a line to tell Minitab to ignore the rest of the line. Text after # is not displayed in the output when the macro is executed (even when you use ECHO).

Array table for OADESIGN

The following table shows all the legal commands and the full arrays to which they correspond.

Command line	Full orthogonal array
OADesign 4 2 K (K=1, 2, 3)	L4 (2**3)
OADesign 8 2 K (K= 1, ... ,7)	L8 (2**7)
OADesign 8 2 K 4 1 or OADesign 8 4 1 2 K (K = 1, 2, 3, 4)	L8 (2**4 4**1)
OADesign 9 3 K (K= 1, 2, 3, 4)	L9 (3**4)
OADesign 12 2 K (K = 1, . . . , 11)	L12 (2**11)
OADesign 16 2 K (K = 1, . . . , 15)	L16 (2**15)
OADesign 16 2 K 4 1 or OADesign 16 4 1 2 K (K = 1, . . . , 12)	L16 (2**12 4**1)
OADesign 16 2 K 4 2 or OADesign 16 4 2 2 K (K = 1, ... , 9)	L16 (2**9 4**2)
OADesign 16 2 K 4 3 or OADesign 16 4 3 2 K (K = 1, . . . , 6)	L16 (2**6 4**3)
OADesign 16 2 K 4 4 or OADesign 16 4 4 2 K (K = 1, 2, 3)	L16 (2**3 4**4)
OADesign 16 4 K (K = 1, . . . , 5)	L16 (4**5)

Command line	Full orthogonal array
OADesign 16 8 1 2 K or OADesign 16 2 K 8 1 (K = 1, . . . , 8)	L16 (8**1 2**8)
OADesign 18 2 1 3 K or OADesign 18 3 K 2 1 (K = 1, . . . , 7)	L18 (2**1 3**7)
OADesign 18 6 1 3 K or OADesign 18 3 K 6 1 (K = 1, . . . , 6)	L18 (6**1 3**6)
OADesign 25 5 K (K = 1, . . . , 6)	L25 (5**6)
OADesign 27 3 K (K = 1, . . . , 13)	L27 (3**13)
OADesign 32 2 K (K = 1, . . . , 31)	L32 (2**31)
OADesign 32 2 1 4 K or OADesign 32 4 K 2 1 (K = 1, . . . , 9)	L32 (2**1 4**9)
OADesign 36 2 K2 3 K3 or OADesign 36 3 K3 2 K2 (K2 = 1, . . . , 11; K3 = 1, . . . , 12)	L36 (2**11 3**12)
OADesign 36 2 K 3 13 (K = 1, 2, 3)	L36 (2**3 3**13)
OADesign 54 2 1 3 K or OADesign 54 3 K 2 1 (K = 1, . . . , 25)	L54 (2**1 3**25)

Assigning attributes to groups

Attributes—such as symbol types, line sizes, and area colors—are assigned to groups in the order the groups appear in the worksheet for text data. For numeric data the groups are assigned from the least to the greatest values.

Assigning attributes with multiple graphs and groups

If you specify several graphs and several groups, then attributes are assigned to the groups and not the graphs.

In the example below, the PLOT command specifies two graphs, which are overlaid (the Advertis label does not appear). The CONNECT subcommand does not specify a group variable, so the two line types are assigned to the two graphs. The SYMBOL subcommand, however, has a group variable, AdAgency, so the two symbol types are assigned to the two ad agencies.

Data

In ResinDefects.MTW: Discoloration Defects, Clump Defects, Hours Since Cleanse, Size of Screw

Dialog box, Step 1: Create the graph

1. Choose **Graph** > **Scatterplot**.
2. Choose **Simple**, then click **OK**.
3. In row 1 of the Data Source, enter *Discoloration Defects* in Y and *Hours Since Cleanse* in X. In row 2, enter *Clump Defects* in Y and *Hours Since Cleanse* in X.
4. Click **Multiple Graphs** and select **Overlaid on the Same Page**. Click **OK**.
5. Click **Data View** and select **Connect Line**.
6. Click **OK** in each dialog box.

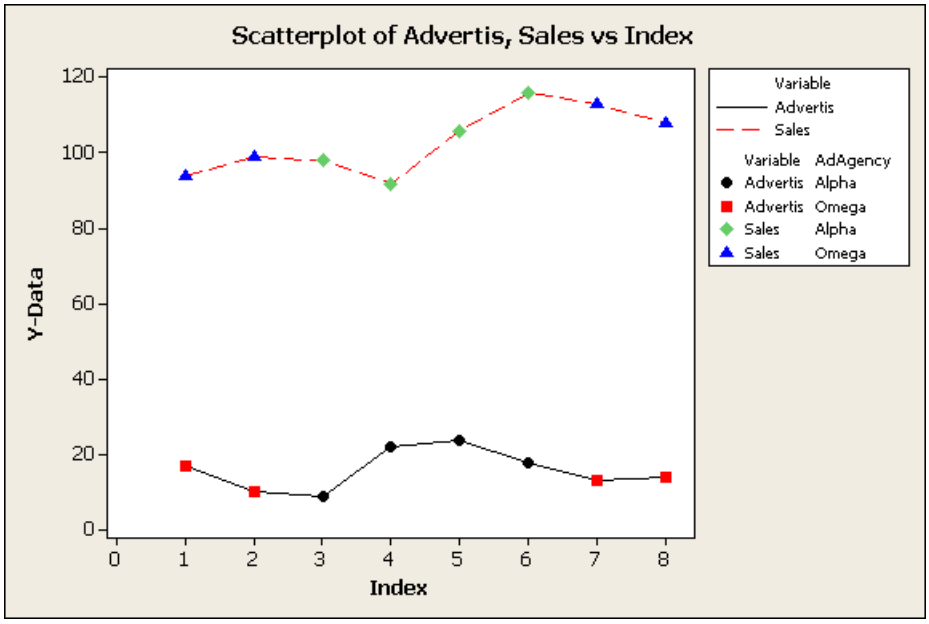
Dialog box, Step 2: Edit the graph

1. Double-click the symbols.
2. On the **Groups** tab, in **Categorical Variable for Grouping**, enter *Size of Screw*.
3. Click **OK**.

Session commands

```
Plot 'Discoloration Defects' * 'Hours Since Cleanse' 'Clump Defects' * 'Hours Since
Cleanse';
Overlay;
NoJitter;
Graph 6 4;
Legend;
Section 1;
Section 2;
Connect;
Symbol 'Size of Screw';
Title;
Footnote;
FPanel;
NoDTitle.
```

Graph output



Bar chart functions

You can create a bar chart of a function of the graph variable using the syntax CHART F(C...C) * C. The following functions (F) are available:

SUM

Sum of values

COUNT
Number of values (frequency)

N
Number of nonmissing values

NMISS
Number of missing values

MEAN
Mean of values

MEDIAN
Median of values

MINIMUM
Minimum of values

MAXIMUM
Maximum of values

PERCENTILE
Use the Kth percentile to summarize the contents of the variable. Must also specify the subcommand `PTILE` K.

PBETWEEN
Use the percentage of values between two values to summarize the contents of the variable. Must also specify the subcommand `PBETWEEN` K1 K2.

PEQUAL
Use the percentage of values equal to one or more values to summarize the contents of the variable. Must also specify the subcommand `PEQUAL` K...K.

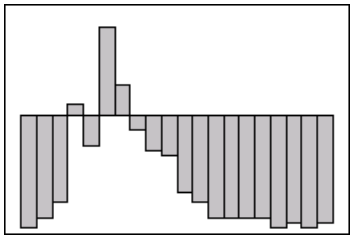
STDEV
Standard deviation of values

SSQ
Sum of squares of values

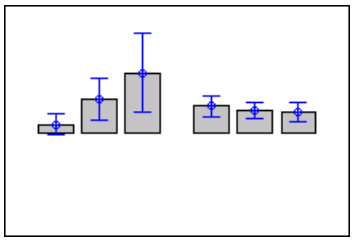
Base position for project lines, area, and bar

You can set the base position for project lines, area, and bar. The default base differs depending on the graph and the data display.

Histogram with non-zero base



Interval plot with bars and non-zero base



Box-Behnken designs

The table below displays a summary of the Box-Behnken designs that are generated by Minitab.

For more information, go to BBDESIGN: Session command for creating a Box-Behnken design on page 343.

Number of factors	Number of runs	Optional number of blocks	Default number of center points
3	15	1	3
4	27	3	3
5	46	2	6
6	54	2	6
7	62	2	6
9	130	5 or 10	10
10	170	2	10

Calculating a chi-square statistic for a goodness-of-fit test using session commands

If you know the expected probability of the possible outcomes, you can calculate a chi-square statistic for a goodness-of-fit test. This topic shows how to do that using session commands.

Equal probability of outcomes

Suppose you ask 20 people at a restaurant whether they prefer bread or rice. You expect that 10 will prefer bread and 10 will prefer rice. You conduct the survey and find that 5 people prefer bread, and 15 people prefer rice. To test whether the observed values differ statistically from the expected values, enter the following commands.

```
SET C1
  5 15 #(observed values)
END
SET C2
 10 10 #(expected values)
END
LET K1 = SUM((C1 - C2)**2/C2) #(chi-square statistic)
PRINT K1
```

K1 contains the correct χ^2 statistic, which is 5. To compute a p-value for that statistic, enter these commands.

```
CDF K1 K2; #(store the CDF in K2)
  CHISQUARE 1. #(df = number of categories - 1)
LET K3 = 1 - K2 #(K3 contains the associated p-value)
PRINT K3
```

The p-value is 0.025, indicating a statistically significant difference between the values we observed and those we expected to find.

Binomial case

Suppose you want to measure how often heads comes up in 5 tosses of a coin, and you repeat those 5 tosses many times. If you toss 5 coins, the number of heads should follow a binomial distribution with n=5 and p=0.5. There are six categories of outcomes: 0 heads in 5 tosses, 1 head in 5 tosses, 2 heads in 5 tosses, up to 5 heads in 5 tosses. You toss coins a total of 3590 times.

```
SET C1
  0 1 2 3 4 5 #(categories)
END
SET C2
  100 524 1080 1126 655 105 #(observed values in each category)
END
PDF C1 C3; #(probability of success for each category)
  BINOMIAL 5 0.5.
LET C4 = 3590 * C3 #(expected values)
LET K1 = SUM((C2 - C4)**2/C4)#(chi-square statistic)
PRINT K1.
CDF K1 K2; #(store the CDF in K2)
  CHISQUARE 5. #(df = number of categories - 1)
LET K3 = 1 - K2 #(K3 contains the associated p-value)
PRINT K3
```

The p-value, stored in K3, is 0.00063, indicating a statistically significant different between what we observed and what we expected.

Calculations for FFACTORIAL

In calculations for FFACTORIAL, Minitab first sets aside any center points and covariates, then checks the design to see if it is orthogonal. If it is not, Minitab displays a message. Next, Minitab checks for collinearity. If there is a problem with factors or blocks, Minitab displays a message and quits. If interaction terms are collinear with other terms in the model, Minitab displays a message, removes the interaction terms, and proceeds. If the design is orthogonal, Minitab also provides information about confounding for the terms it removes. Minitab then adds in any center points and covariates and does a final collinearity check.

FFACTORIAL uses a regression approach to fit the model. This approach allows you to analyze nonorthogonal designs (also called unbalanced designs) as well as orthogonal designs. However, if the design is large, especially if there are many terms, the calculations might take a long time.

FFACTORIAL can fit models with up to 137 terms and up to 15-way interactions. However, on computers that have limited worksheet space, Minitab might not be able to fit all 137 terms. For example, on a computer that has a worksheet of 16,000, Minitab can fit only 64 terms.

Central composite designs

The following table is a summary of the designs that are created by Minitab's central composite design commands.

Factors	Design	Total runs	Total blocks	Cube blocks	Cube runs	Total center points	Cube center points	Axial center points	Default alpha	Orthogonal blocks?	Rotatable?
2	Full	13	1	---	4	5	---	0	1.414	---	Yes
2	Full	14	2	1	4	6	3	3	1.414	Yes	Yes
3	Full	14	1	---	8	6	---	0	1.682	---	Yes
3	Full	20	2	1	8	6	4	2	1.683	Yes	No
3	Full	20	3	2	8	6	4	2	1.683	Yes	No
4	Full	31	1	---	16	7	---	0	2.000	---	Yes
4	Full	30	2	1	16	6	4	2	2.000	Yes	Yes
4	Full	30	3	2	16	6	4	2	2.000	Yes	Yes
5	Full	52	1	---	32	10	---	0	2.378	---	Yes
5	Full	54	2	1	32	12	8	4	2.366	Yes	No
5	Full	54	3	2	32	12	8	4	2.366	Yes	No
5	Half	32	1		16	6	---	0	2.000		Yes
5	Half	33	2	1	16	7	6	1	2.000	Yes	Yes
6	Full	90	1	---	64	14	---	0	2.828	---	Yes
6	Full	90	2	1	64	14	8	6	2.828	Yes	Yes
6	Full	90	3	2	64	14	8	6	2.828	Yes	Yes
6	Full	90	5	4	64	14	8	6	2.828	Yes	Yes
6	Half	53	1	---	32	9	---	0	2.378	---	Yes
6	Half	54	2	1	32	10	8	2	2.366	Yes	No
6	Half	54	3	2	32	10	8	2	2.366	Yes	No
7	Full	152	1	---	128	1	---	0	3.364	---	Yes
7	Full	160	2	1	128	18	8	10	3.364	Yes	Yes
7	Full	160	3	2	128	18	8	10	3.364	Yes	Yes
7	Full	160	5	4	128	18	8	10	3.364	Yes	Yes
7	Half	88	1	---	64	10	---	0	2.828	---	Yes
7	Half	90	2	1	64	12	8	4	2.828	Yes	Yes
7	Half	90	3	2	64	12	8	4	2.828	Yes	Yes
7	Half	90	5	4	64	12	8	4	2.828	Yes	Yes
8	Half	154	1	---	128	10	---	0	3.364	---	Yes
8	Half	160	2	1	128	16	8	8	3.364	Yes	Yes
8	Half	160	3	2	128	16	8	8	3.364	Yes	Yes
8	Half	160	5	4	128	16	8	8	3.364	Yes	Yes
8	Quarter	90	1	---	64	10	---	0	2.828	---	Yes
8	Quarter	90	2	1	64	10	8	2	2.828	Yes	Yes
8	Quarter	90	3	2	64	10	8	2	2.828	Yes	Yes

Factors	Design	Total runs	Total blocks	Cube blocks	Cube runs	Total center points	Cube center points	Axial center points	Default alpha	Orthogonal blocks?	Rotatable?
8	Quarter	90	5	4	64	10	8	2	2.828	Yes	Yes
9	Quarter	156	1	---	128	10	---	0	3.364	---	Yes
9	Quarter	160	2	1	128	14	8	6	3.364	Yes	Yes
9	Quarter	160	3	2	128	14	8	6	3.364	Yes	Yes
9	Quarter	160	5	4	128	14	8	6	3.364	Yes	Yes
10	Eighth	158	1	---	128	10	---	0	3.364	Yes	Yes
10	Eighth	160	2	1	128	12	8	4	3.364	Yes	Yes
10	Eighth	160	3	2	128	12	8	4	3.364	Yes	Yes
10	Eighth	160	5	4	128	12	8	4	3.364	Yes	Yes

Computing weights, or smoothed values

The smoothed values are obtained in one of two ways:

- Using the ARIMA (0,1,1) model (without WEIGHT subcommand). In this method, the ARIMA model is fit to the data using Minitab's ARIMA command. The predicted values from the ARIMA command are stored. These are the same as the smoothed values, only lagged one period. There is no predicted value for period 1, so this must be obtained. This is done by backcasting period 1. Thus, predicted in period 1 = [predicted in period 2 – a(data in period 1)] / (1 – a).

- Using the classical approach (with WEIGHT subcommand). In this method, the average of the first 6 (or N, if N < 6) data values is used to obtain the initial smoothed value. Once the initial smoothed value is obtained, subsequent values are calculated from the formula: smoothed at period t = (a) * (data at period t) + (1 – a) * (smoothed at period t – 1).

Data requirements and formats for attribute agreement analysis

For attribute agreement analysis, data can be text or numeric. The assigned ratings can be nominal or ordinal:

- Nominal data are categorical variables that have two or more levels with no natural ordering. For example, the levels in a food tasting study may include crunchy, mushy, and crispy.
- Ordinal data are categorical variables that have three or more levels with a natural ordering, such as strongly disagree, disagree, neutral, agree, and strongly agree.

If the response is a text column and the data are ordinal, click in the column, then right-click and choose **Column Properties** > **Value Order** to define a non-alphabetical ordering on the text values.

Data can be stacked in an attribute column or unstacked in multiple columns. For attribute column data, stack all of the responses in one column and set up grouping indicator columns for the appraiser and part number. The grouping indicators define each sample. Grouping indicators and responses can be text or numeric. For multiple column data, responses for each appraiser and each trial are in separate columns.

The following worksheets contain the same data set, structured in both of the formats that you can use for attribute agreement analysis.

Attribute column data

C1	C2	C3
Sample	Appraiser	Response
1	A	Good
1	A	Good
1	B	Bad
1	B	Good
2	A	Good
2	A	Good
2	B	Good
2	B	Good
3	A	Bad
3	A	Good
3	B	Bad
3	B	Bad
4	A	Good
4	A	Good
4	B	Good
4	B	Good
5	A	Bad
5	A	Bad
5	B	Good
5	B	Bad

Multiple columns of data

C1	C2	C3	C4	C5
Sample	Appraiser A, Trial 1	Appraiser A, Trial 2	Appraiser B, Trial 1	Appraiser B, Trial 2
1	Good	Good	Bad	Good
2	Good	Good	Good	Good
3	Bad	Good	Bad	Bad
4	Good	Good	Good	Good
5	Bad	Bad	Good	Bad

Note When your data are in multiple columns, you must enter the trials for each operator together, as shown above. The order of the trials must be the same for each operator.

Default date/time formats

Minitab recognizes many date/time formats, including the following.

Format	Description	Example
d-mmmm-yy	International date	5-January-16
mmmm d	Month and day	January 5
d mmmm	Day and month	5 January
d mmmm yy	Day, month, and year	5 January 16
mmmm d, yy	Month, day, and year	January 5, 16
dddd, mmmm dd, yyyy	Day of the week, month, date, and year	Sunday, January 05, 2014
h:mm.ss.ss AM/PM	12-hour time	11:10:14.22 PM
[m]:ss.sss	Elapsed time (minutes, seconds, thousandths of a second)	2193:44.123
[h]:mm:ss	Elapsed time (hours, minutes, seconds)	5:23:14
[d].hh:mm	Elapsed time (days, hours, minutes)	1.12:33

Minitab also recognizes variations of these formats. For example, Minitab recognizes mmmm d, yy (January 1, 14) and also recognizes mmm dd, yyyy (Jan 01, 2014).

If you enter time components without a date or without AM/PM, such as 1:17, Minitab recognizes the value as an elapsed time (one hour and 17 minutes) not as a specific time on the 12- or 24-hour clock.

Date/time formats from regional settings

Minitab recognizes the short date format, the long date format, the short time format, and the long time format from your regional date and time settings on your computer. Minitab also recognizes date/time formats based on these short and long formats. For example, regionally-based default formats for January 12, 2014 include.

Format	Regional setting	Example
m/d/yy	English (United States)	1/12/16
yy/m/d	Chinese, Korean, or Japanese	16/1/12
d/m/yy	English (United Kingdom) or French (France)	12/1/16

Minitab interprets 2-part dates (month and day, without the year) based on regional settings. For example, if the regional setting is m/d/yy or yy/m/d, Minitab interprets 1/12 as January 12 of the current year. However, if the regional setting is yy/d/m, Minitab interprets the 2-part date 1/12 as December 1 of the current year. Minitab always interprets a 2-part date that includes a 4-digit year as a month and a year.

Separators and symbols

Default date components can be separated by (/), (-), or (.). All components must be separated by the same symbol (except for the default format mmmm d, yy). For example, both 1/1/99 and 1-1-99 are valid, but 1/1-99 is not valid.

Default 12-hour time components can be AM/PM, am/pm, A/P, or a/p.

Square brackets are used to indicate elapsed time, with the brackets around the most significant time unit. For example, [m]:ss represents elapsed time in minutes and seconds. If you enter time components without a date or without AM/PM (such as 1:17) into the worksheet, Minitab recognizes the value as an elapsed time (one hour and 17 minutes) not as

a specific time on the 12- or 24-hour clock. The format for fraction of a second can be .s (tenths), .ss (hundredths), or .sss (thousandths).

Note When the regional language is set to German, French, or Portuguese, the default decimal separator is a comma. For example, in English a fraction of a second uses a period (ss.sss). But in the three languages previously mentioned, a fraction of a second uses a comma (ss,sss).

Variations of default formats

Minitab also recognizes variations of the above default formats. For example, the time components h:mm:ss can be added to a default date format to create a longer date-time format, such as m/d/yyyy h:mm:ss. In addition, some date/time components (such as "d" or "mmmm") can be substituted with longer or shorter forms ("dd" or "mmm").

Date component	What the component can be substituted with
d	dd
m	mm
yy	yyyy
mmmm	mmmm
h	hh
s	ss

For example, Minitab recognizes mmmm d, yy (January 1, 16 but also recognizes mmm dd, yyyy (Jan 01, 2016).

Note In Korean and Japanese, mmm cannot be substituted for mmmm in a default format. In these languages, the mmm format for month is a number only, which may create ambiguity with other default formats that use only a number to represent the month (mm and m).

Designs generated by FFDESIGN

The table below summarizes the designs generated by FFDESIGN. The number in a cell is the resolution of the unblocked design.

	Number of factors													
Number of runs	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
4	full 2	III 1												
8		full 4	IV 4	III 2	III 2	III 1								
16			full 8	V 8	IV 8	IV 8	IV 8	III 4	III 4	III 4	III 4	III 2	III 2	III 1
32				full 16	VI 16	IV 8	IV 8	IV 8	IV 8	IV 8	IV 8	IV 8	IV 8	IV 8
64					full 32	VII 16	V 16	IV 16	IV 16	IV 16	IV 16	IV 16	IV 16	IV 16
128						full 64	VIII 32	VI 16	V 16	V 16	IV 16	IV 16	IV 16	IV 16

Entering data for factor variables

You can enter numeric or text data to assign values to factor levels. For a numeric factor, Minitab designates the lowest numeric value as the reference level. For a text factor, Minitab determines the reference level alphabetically.

Logistic regression creates a set of design variables for each factor in the model. Minitab uses the reference cell method to code design variables that make up a factor. For each factor, Minitab creates the number of levels minus one design variables.

Here are two examples of the default coding scheme.

Factor A with 4 levels

The reference level is 1.

	A1	A2	A3
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1

Factor B with 3 levels

The reference level is Humidity

	B1	B2
Humidity	0	0
Pressure	1	0
Temperature	0	1

Use the REFERENCE subcommand to specify the reference level for a factor. Use the EFFECT subcommand to specify effect coding instead of binary coding.

Entering data for response variables

The response variable data can be represented in various ways. Minitab provides four response variable subcommands—FREQUENCY, ST, SF, and FT—to specify the form of the response data. The table below summarizes the forms.

For examples, go to Examples of entering response data for logistic regression on page 1227.

How the response data is entered	Description	Commands	Response variable subcommand
One column that contains distinct values	Response column contains one row for each observation, with additional column or columns to specify the factor/covariate patterns	GZLM OLOGISTIC NLOGISTIC	None
Frequencies	One column contains the response values, one column contains the frequencies associated with the response, with additional column or columns	GZLM OLOGISTIC	FREQUENCY

How the response data is entered	Description	Commands	Response variable subcommand
	to specify the factor/covariate patterns for each frequency	NLOGISTIC	
Events and trials or trials	Two columns represent the response variable with additional column or columns to specify the factor/covariate pattern for each observation	GZLM BFFACTORIAL BGFACTORIAL BRSREGRESSION BSCREEN	None

Entering patterned data for the SET session command

SET: Session command for entering data into a column on page 81 provides shortcut ways to enter data that follow many different kinds of patterns.

The following example shows how to set C3 with the numbers 1, 2, 3, 4, and 5. Note that a colon abbreviates the range.

```
SET C3
 1:5
 END
```

Observe the following rules:

1. A slash (/) indicates an increment.
2. Put lists you want repeated in parentheses.
3. No space between the repeat factor and the corresponding parenthesis.
4. If the repeat factor is before the parentheses, Minitab repeats the entire list; if it is after, Minitab repeats each number in the range.
5. You cannot nest parentheses.
6. You can use a stored constant in place of any number. You can indicate a series of stored constants on the data line with a "-" (K1-K5).

   4:1 expands to 4, 3, 2, 1

   1:3/.5 expands to 1, 1.5, 2, 2.5, 3

   3(1) expands to 1, 1, 1

   3(1:3) expands to 1, 2, 3, 1, 2, 3, 1, 2, 3

   (1:3)2 expands to 1, 1, 2, 2, 3, 3

   3(1:3)2 expands to 1, 1, 2, 2, 3, 3, 1, 1, 2, 2, 3, 3, 1, 1, 2, 2, 3, 3

Restricted and unrestricted mixed models

A company ran an experiment to see how several conditions affect the thickness of a coating substance that it manufactures. The experiment was run at two different times, in the morning and in the afternoon. Three operators were chosen from a large pool of operators employed by the company. The manufacturing process was run at three settings, 35, 44, and 52. Two determinations of thickness were made by each operator at each time and setting. Thus, the three factors are crossed. One factor, operator, is random. The other factors, time and setting, are fixed.

The statistical model is as follows:

$Y_{ijkl} = \mu + T_i + O_j + S_k + TO_{ij} + TS_{ik} + OS_{jk} + TOS_{ijk} + e_{ijkl}$

where T_i is the time effect, O_j is the operator effect, and S_k is the setting effect, and TO_{ij}, TS_{ik}, OS_{jk}, and TOS_{ijk} are the interaction effects.

Operator, all interactions with operator, and error are random. The random terms are as follows:

O_j TO_{ij} OS_{jk} TOS_{ijk} e_{ijkl}

These terms are all assumed to be normally distributed random variables with mean zero and variances given by the following:

var $(O_j) = V(O)$

var $(TOS_{jkl}) = V(TOS)$

var $(TO_{ij}) = V(TO)$

var $(e_{ijkl}) = V(e) = \sigma^2$

These variances are called variance components. The output from expected means squares contains estimates of these variances.

In the unrestricted model, all these random variables are independent. The remaining terms in this model are fixed.

In the restricted model, any term which contains one or more subscripts corresponding to fixed factors is required to sum to zero over each fixed subscript. In the example, this means:

$$\sum_i (T_j) = 0 \qquad \sum_k (S_k) = 0 \qquad \sum_i (TO_{ij}) = 0$$

$$\sum_k (TS_{ik}) = 0 \qquad \sum_k (OS_{jk}) = 0 \qquad \sum_i (TOS_{ijk}) = 0$$

Your choice of model does not affect the sums of squares, degrees of freedom, mean squares, or marginal and cell means. It does affect the expected mean squares, error term for the F-tests, and the estimated variance components.

Examples of entering response data for logistic regression

When the response variable is in one column that contains distinct values

When you enter the response as raw data, you do not need a response variable subcommand. The response column will contain a limited number of distinct values—2 for binary responses, and 3 or more distinct values for ordinal and nominal responses. The examples below all use the same binary response variable.

C1	C2	C3	C4
Response		Factor	Covar
0		1	12
1		1	12
1		1	12
1		1	12
0		2	12
1		2	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		1	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12

C1	C2	C3	C4
Response		Factor	Covar
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
1		2	12
0		1	24
0		1	24
0		1	24
0		1	24
0		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24
1		1	24

In this form, the response variable is in C1, the factor is in C3, and the covariate is in C4. To fit the full model, type:

```
GZLM;
 BINOMIAL;
 LOGIT;
 RESPONSE C1;
 CONTINUOUS C4;
 CATEGORICAL C3;
 TERMS C3|C4.
```

When the response variable is entered as frequencies

You can also represent the response data with one column of frequencies. For each row of data, the frequencies are the number of times the response/factor/covariate pattern occurred. For example, the data used above could have been entered in the following manner.

C1	C2	C3	C4
Response	Count	Factor	Covar
0	1	1	12
1	19	1	12
0	1	2	12
1	19	2	12
0	5	1	24
1	15	1	24
0	4	2	24
1	16	2	24
0	7	1	50
1	13	1	50
0	8	2	50
1	12	2	50
0	11	1	125
1	2	1	125
0	9	2	125
1	11	2	125
0	19	1	200
1	1	1	200
0	18	2	200
1	2	2	200

This form represents the same binary data as above, however, the frequencies for each response/factor/covariate pattern are in C2. To fit the full model, type the following:

```
GZLM;
 BINOMIAL;
 LOGIT;
 RESPONSE C1;
 FREQUENCY C2;
 CONTINUOUS C4;
 CATEGORICAL C3;
 TERMS C3|C4.
```

When the response variable is entered as the number of events and trials

You can also enter the response variable as the number of events and trials. This method is available when you have a binary response variable. GZLM, BFFACTORIAL, BRSREGRESS, BGFACTORIAL, and BSCREEN are examples of commands that analyze binary response variables.

C1	C2	C3	C4
E	T	Factor	Covar
19	20	1	12
19	20	2	12

C1	C2	C3	C4
E	T	Factor	Covar
15	20	1	24
16	20	2	24
13	20	1	50
12	20	2	50
9	20	1	125
11	20	2	125
1	20	1	200
2	20	2	200

In this form, the data are represented by the number of successes in C1, and the number of trials in C2. To fit the full model, type the following

```
GZLM;
 BINOMIAL;
 LOGIT;
 RESPONSE C1 C2;
 CONTINUOUS C4;
 CATEGORICAL C3;
 TERMS C3|C4.
```

A comparison of MSURFACE and RMCONTOUR plots

Three-dimensional surface plots are oriented differently than response surface contour plots. You can generate a three-dimensional response surface plot so that the view is similar to the associated response surface contour plot, as follows:

1. Switch the x- and y-axes means with the SINGLE subcommand. Type the y-variable and the x-variable factors respectively, as arguments.
2. Change where you view the plot from by changing the view position. Use the VPOSITION subcommand with −5 −4 3 as arguments.

Graphics options for MIXCONTOUR

MIXCONTOUR on page 393 supports the following graphics subcommands (and a few commands). Many of the subcommands have subcommands of their own.

Scale

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072: When you use AREA with MIXCONTOUR, you cannot use the following subcommands: BASE, ETYPE, ECOLOR, ESIZE, DIRECTION, STRAIGHT, and STEP.

CONNECT: Session subcommand for connecting points with lines on page 1085: With CONNECT, the subcommands ORDER, STRAIGHT, and STEP cannot be used.

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot on page 1152

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options for MIXOVER

MIXOVER on page 401 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Click the subcommand or command in the table for more information.

Scale

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options for MIXSURFACE

MIXSURFACE: Session command for creating a surface plot on page 398 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Choose a subcommand or command in the table below for more information and examples.

Options

LBRIGHT: Session subcommand for specifying the brightness of the lights that illuminate a surface plot on page 1117

LIGHT: Session subcommand for specifying the position, color, and visibility of the lights that illuminate a surface plot on page 1120

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CONNECT: Session subcommand for connecting points with lines on page 1085

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot on page 1152

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

VASPECT: Session subcommand for specifying the aspect ratio of the data box on page 1164

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph on page 1165

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units on page 1165

VPOSITION: Session subcommand for specifying the view position as a ratio on page 1165

VUP: Session subcommand for specifying which direction is up in relation to the data box on page 1165

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options behavior that is specific to MIXSURFACE

The following is a summary of changes in the behavior of common graphics options with MIXSURFACE:

- SYMBOL does not support symbol label functionality for three-dimensional surface plots.
- The defaults for these subcommands vary depending on if you use menu commands or use MIXSURFACE from the command line. These defaults are described in the subcommand descriptions.

With AXLABEL, the following subcommands are unusable: LABEL, LSHOW, and SECSCALE.

With CONNECT, the following subcommands are unusable: ORDER, STRAIGHT, and STEP.

Graphics options for MSURFACE

MSURFACE on page 1181 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Choose a subcommand or command in the table below for more information and examples.

Scale

MGRID, NOGRID, and NOMGRID: Session subcommands for controlling the grid on a graph on page 1102

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

CONNECT: Session subcommand for connecting points with lines on page 1085

SFIT: Session subcommand for specifying the attributes of the surface on a 3D surface plot on page 1152

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

VASPECT: Session subcommand for specifying the aspect ratio of the data box on page 1164

VBOX: Session subcommand for specifying the display of the box that surrounds the data on a graph on page 1165

VFIELD: Session subcommand for specifying the dimensions of the field of view in object units on page 1165

VPOSITION: Session subcommand for specifying the view position as a ratio on page 1165

VUP: Session subcommand for specifying which direction is up in relation to the data box on page 1165

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options behavior that is specific to MSURFACE

This is a summary of changes in the behavior of common graphics options with MSURFACE:

- SYMBOL does not support symbol label functionality for three-dimensional surface plots.
- The defaults for these subcommands vary depending on if you use menu commands or use MSURFACE from the command line. These defaults are described in the subcommand descriptions.

With AXLABEL, LABEL, LSHOW, and SECSCALE cannot be used.

With CONNECT, the subcommands ORDER, STRAIGHT, and STEP cannot be used.

Graphics options for SIMPLEX

SIMPLEX on page 385 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Click the subcommand or command in the table for more information.

Scale

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options for SPCONT

SPCONT on page 395 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Click the subcommand or command in the table for more information.

Scale

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Data view

AREA: The session subcommand for shading the area below the data values to the base on page 1072

CONNECT: Session subcommand for connecting points with lines on page 1085

SYMBOL: Session subcommand for displaying a symbol for each data value on page 1157

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphics options behavior that is specific to SPCONT

This is a summary of changes in the behavior of common graphics options with SPCONT. This information also appears under the respective graphics options.

With AREA, the subcommands BASE, ETYPE, ECOLOR, ESIZE, DIRECTION, STRAIGHT, and STEP cannot be used.

With CONNECT, the subcommands ORDER, STRAIGHT, and STEP cannot be used.

Graphics options for SPSIMP

SPSIMP on page 387 supports the graphics subcommands (and a few commands) in the following table. The subcommands are shown without syntax. Many of these subcommands have subcommands of their own. Click the subcommand or command in the table for more information.

Scale

NOFRAME: Session subcommand for suppressing lines and labels on a graph on page 1130

REFERENCE: Session subcommand for specifying the axis and location of reference lines on a graph on page 1146

SCALE: Session subcommand for customizing the axes and ticks of a graph on page 1149

TRANSPOSE and NOTRANSPOSE: Session subcommands for transposing the x- and y-axis on a graph on page 1164

Labels

AXLABEL: Session subcommand for customizing graph axis labels on page 1074

FOOTNOTE: Session subcommand for adding a footnote to a graph on page 1098

NODTITLE, NODSUBTITLE, and NODFOOTNOTE: Session subcommands for suppressing titles, subtitles, and footnotes on a graph on page 1130

NOPERFOOTNOTE: Session subcommand for suppressing My Footnote on page 1131

SUBTITLE: Session subcommand for adding a subtitle to a graph on page 1156

TITLE: Session subcommand for adding a title to a graph on page 1162

Multiple graphs

OVERLAY: Session subcommand for combining graphs specified in a multiple graph command into a single graph on page 1134

PANEL: Session subcommand for paneling graphs on page 1134

SAME: Session subcommand for specifying that one or more axes are the same for multiple graphs on page 1149

SEPARATE: Session subcommand for placing groups in separate graphs when you use a paneling variable on page 1152

Data options

NOEMPTY and NOMISS: Session subcommands for excluding missing data from graphs on page 1130

Annotation

ELLIPSE: Session subcommand for constructing an ellipse from points on a graph on page 1095

LINE: The session subcommand for constructing a line from points on a graph on page 1121

MARKER: Session subcommand for displaying a symbol at specified points on a graph on page 1123

POLYGON: Session subcommand for constructing a polygon from points on a graph on page 1136

RECTANGLE: Session subcommand for constructing a rectangle from points on a graph on page 1145

TEXT: Session subcommand for displaying text on a graph on page 1161

Region

FIGURE: Session subcommand for controlling the figure region within the graph region on page 1096

GRAPH: Session subcommand for controlling the graph region fill and border line on page 1101

LEGEND and NOLEGEND: Session subcommands for controlling the legend on a graph on page 1118

NOHLEGEND: Session subcommand for suppressing the legend of hold values on a graph on page 1131

Graph, input, and output

GSAVE: Session subcommand for saving a graph in a file on page 39

WTITLE: Session subcommand for specifying the title of the output pane on page 1166

Graphs that use groups with the data display subcommands

PLOT: Session command for creating a scatterplot on page 1020

BOXPLOT: Session command for creating a boxplot on page 1048

MATRIXPLOT: Session command for creating a matrix of plots on page 1025

TSPLOT: Session command for creating a time series plot on page 1060

CHART: Session command for creating a bar chart on page 1012

HISTOGRAM: Session command for creating a histogram on page 1031

PLTX: Session command for creating a 3D scatterplot on page 1068

SURFACEPLOT: Session command for creating a surface plot on page 1069

The control charts and CONTOURPLOT: Session command for creating a contour plot on page 1065 do not use grouping variables with the Data Display subcommands. %graphs do not use data display subcommands.

How to enter data for ANOVA and GLM

The data set has the same structure in ANOVA on page 225 and GLM on page 228.

You need one column for each response variable and one column for each factor. This means there is one row of the worksheet for each observation. This row contains the value of each response variable and the level of each factor.

Regardless of whether factors are crossed or nested, you use the same form for the data. Factor columns may be numeric or text, and may contain any values. The levels do not need to be in any special order.

There is a restriction on the subscripts used for nested terms as nesting must be balanced. Suppose A has 3 levels, and B is nested within A. If B has 4 levels within the first level of A, B must have 4 levels within the second and third levels of A. Minitab will tell you if you have unbalanced nesting. In addition, the subscripts used to indicate the 4 levels of B within each level of A must be the same. Thus, the four levels of B cannot be (1 2 3 4) in level 1 of A, (5 6 7 8) in level 2 of A, and (9 10 11 12) in level 3 of A.

The patterned data capability of the SET on page 81 command can be helpful in entering the factor levels. For example, to enter the level numbers for a three-way crossed design with 3 levels of 3 factors and 1 observations per cell, use:

```
set c1
(1:3)9
end.
set c2
3(1:3)3
end
set c3
9(1:3)1
end
```

How to enter data for CCHART

You must have at least one column containing the number of defects. Each row of a column should contain the number of defects for one subgroup.

Suppose you collect data on the number of scratches found per square meter of finished surface each day. You sample the same number of square meters each day. The number of defects is the number of scratches found on a given day. Your data might like the Scratches column below.

C1
Scratches
2
3
1
3
0
4

When an observation is missing, a gap exists in the chart for that observation.

How to enter data for GCHART

To create a G chart, you can record your data in one of the following ways:

- Dates of events
- Number of opportunities between events
- Number of opportunities until an event

Dates of events

When your data are entered as dates, Minitab calculates the number of days between events and plots each value on the G chart. When you use dates, you should verify that the number of possible opportunities per day is fairly constant.

C1
Date of infection
4/17/16

C1
Date of infection
4/24/15
4/26/16
5/14/16
5/16/16

Number of opportunities between or until an event

You can record either the number of opportunities between events or the number of opportunities until an event occurs.

Suppose infections occurred in the 4^{th}, 16^{th}, 27^{th}, 50^{th}, and 80^{th} patient in an emergency room. If you record the number of patients between events, the data are shown below in the Between column. If you record the number of opportunities until an event, the data are shown below in the Until column.

C1	C2
Between	Until
11	12
10	11
22	23
29	30

If your data are recorded as the number of opportunities until an event, Minitab subtracts 1 from each value to convert the data to number of opportunities between events and plots the converted values on the G chart.

How to enter data for NPCHART

You must have at least one column containing the number of defectives. Each row of a column should contain the number of defectives for one subgroup. If the subgroup sizes are different, you must also include a column that indicates the size of each subgroup. If the size is the same for each subgroup, you can enter the subgroup size directly at the command line instead of providing a column.

Suppose that, at a manufacturing plant, you collect data each day on the number of parts that failed inspection. The number of defectives is the number of parts that failed on a given day. The subgroup size is the number of parts that you inspected that day. Your data might look like the table below.

Each row in the table represents a single day. The Failed column indicates the number of defectives for the day. The Inspected column indicates the number of parts inspected that day.

C1	C2
Failed	Inspected
2	38
3	56
1	44
3	41
0	54

C1	C2
Failed	Inspected
4	35

Subgroup sizes that differ affect both the control limits and the center line. In general, the control limits are further from the center line for smaller subgroups than they are for larger ones. You can force the control limits and center line to be constant by entering a fixed subgroup size, for example the average subgroup size.

When an observation is missing, a gap exists in the chart for that observation.

How to enter data for PDIAGNOSTIC, PCHART, and PPRIMECHART

You must have at least one column that contains the number of defectives. Each row of a column should contain the number of defectives for one subgroup.

If the subgroup sizes are different, you must also include a column that indicates the size of each subgroup. If the size is the same for each subgroup, you can enter the subgroup size directly at the command line instead of providing a column.

Suppose that, at a manufacturing plant, you collect data each day on the number of parts that failed inspection. The number of defectives is the number of parts that failed on a given day. The subgroup size is the number of parts that you inspected that day. Your data might look like the table below.

Each row in the table represents a single day. The Failed column indicates the number of defectives for the day. The Inspected column indicates the number of parts inspected that day.

C1	C2
Failed	Inspected
2	114
0	108
3	133
2	80
1	109
1	99

The P Chart and Laney P' Chart divide the number of defectives by the subgroup size to calculate the proportion of defectives for each subgroup.

The control limits for the P Chart and Laney P' Chart change depending on the size of each subgroup. In general, the control limits are further from the center line for smaller subgroups than they are for larger ones. You can force the control limits to be constant by entering a fixed subgroup size, for example the average subgroup size.

When an observation is missing, a gap exists in the chart for that observation.

How to enter data for TCHART

To create a T chart, you can record your data in one of the following ways:

- Dates and times of events
- Time between events
- Number of intervals between events

Dates and times of events

When your data are entered as dates and times, Minitab calculates the amount of time between events and plots each value on the T chart.

C1
Date and time of patient fall
1/22/2016 8:05AM
1/25/2016 4:30PM
2/3/2016 5:45AM
2/5/2016 11:30AM
2/9/2016 9:30PM

Time between events

Time between events, also known as elapsed time, must be entered in the worksheet as hours, minutes, and seconds (hh:mm:ss format). If you do not record seconds, Minitab reads the seconds as 00. For example, Minitab reads the first value below as 80:25:00.

C1
Time between patient falls
80:25
205:15
53:45
106:00

Number of intervals between events

You can record the number of intervals between events as days or hours. The number of days or hours must be greater than or equal to 0. A value of 0 indicates that two events happened simultaneously.

C1
Number of days between patient falls
3.33
8.5
2
4.5

How to enter data for UDIAGNOSTIC, UCHART, and UPRIMECHART

You must have at least one column containing the number of defects. Each row of a column should contain the number of defects for one subgroup.

If the subgroup sizes are different, you must also include a column that indicates the size of each subgroup. If the size is the same for each subgroup, you can enter the subgroup size directly at the command line instead of providing a column.

Suppose you collect data each day on the number of scratches found per square meter of finished surface. The number of defects is the number of scratches found on a given day. The subgroup size is the number of square meters of surface that was inspected that day. Your data might look like the table below.

Each row in the table represents a single day. The Scratches column indicates the number of defects found that day. The Square Meters column indicates the number of square meters of surface that was inspected that day.

C1	C2
Scratches	**Square Meters**
8	68
3	54
11	71
8	67
9	73
1	56

The U chart and Laney U' chart divide the number of defects by the subgroup size to calculate the number of defects per subgroup.

The control limits for the U chart and Laney U' chart change depending on the size of each subgroup. In general, the control limits are further from the center line for smaller subgroups than they are for larger ones. You can force the control limits and center line to be constant by entering a fixed subgroup size, for example the average subgroup size.

When an observation is missing, a gap exists in the chart for that observation.

How to enter subgroup data

You can structure your subgroup data in the worksheet down a column or across rows. The worksheets below contain the same data, with subgroups of 5, structured both ways.

Subgroup data that is entered down a column

C1
40.13
39.68
40.84
41.52
40.02

C1
39.54
39.87
40.25
40.47
41.41
39.20
38.72
39.39

Subgroup data that is entered across rows

C1	C2	C3	C4	C5
40.13	39.68	40.84	41.52	40.02
39.54	39.87	40.25	40.47	41.41
39.64	37.36	39.20	38.72	39.39
39.15	41.15	39.96	40.48	40.05

Subgroups in rows

Use the RSUB subcommand to specify the data columns. The following command language specifies that the data is in subgroups of size 5, placed in rows across five columns.

```
XBAR;
  RSUB C1-C5.
```

Subgroups or individual observations in columns: Individual observations

For commands that allow only a data column as an argument, specify the main command and the data column.

```
ICHART C1
```

Some commands that use this format include ICHART, MRCHART, IMRCHART, NNCAPA, and NNSIXPACK.
To enter the data for commands that analyze individuals data or subgroup data, specify the main command with the data column and the size of the subgroup.

```
EWMACHART C1 1
```

Subgroups or individual observations in columns: Same-size subgroups

To enter the data, specify the main command with the data column and the size of the subgroup. The following command language specifies to take the first 5 observations in C1 to be the first subgroup, the second 5 observations to be the second subgroup, and so on.

```
XBARCHART C1 5
```

Some commands that use this format include BOXCOX, BWCHART, BWCAPA, BWSIXPACK, CAPA, CUSUM, EWMACHARRT, GENVAR, MACHART, MEWMA, RCHART, RRUN, SCHART, SIXPACK, TSGV, TSQUARED, XBARCHART, XRCHART, XSCHART, and ZONE.

Subgroups or individual observations in columns: Unequal-size subgroups

When subgroups are of unequal size, you must enter your data in one column, then create a second column of subscripts which serve as subgroup indicators. In the following example, C1 contains the measurement data and C2 contains subgroup indicators.

C1	C2
39.68	1
40.84	1
41.52	1
40.02	2
39.54	2
39.87	2
40.25	2
40.47	2
41.41	2
39.20	3
38.72	3
39.39	3

Each time a subscript changes in C2, a new subgroup begins in C1. In this example, subgroup 1 has 3 observations, subgroup 2 has 6 observations, and so on.

To enter the data, specify the main command with the data column and subgroup column.

```
XBARCHART C1 C2
```

Some commands that use this format include BOXCOX, BWCHART, BWCAPA, BWSIXPACK, CAPA, EWMACHART, GENVAR, MACHART, MEWMA, RCHART, SCHART, SIXPACK, TSGV, TSQUARED, XBARCHART, CRCHART, XSCHART, and ZONE.

How to specify the model for ATCLASS, GZLM, OLOGISTIC and NLOGISTIC

OLOGISTIC on page 214 and NLOGISTIC on page 217 can fit models with the following:

- Up to nine factors and up to 50 covariates
- A binary, ordinal, or nominal response variable
- Crossed or nested factors
- Covariates that are crossed with each other or with factors, or nested within factors

ATCLASS on page 897 and GZLM on page 200 can fit models with any number of factors and responses, but do not model nested terms.

Model continuous predictors as covariates, and categorical predictors as factors. The following are some examples. A is a factor, X is a covariate, and Y is the response.

Command language	Description

`GZLM;`
`RESPONSE Y;`
`CONTINUOUS X;`
`CATEGORICAL A;`
`TERMS A X A*X.`
	Fits a full model with a covariate crossed with a factor

`OLOGISTIC Y = A X X*X;` `FACTORS A.`	Fits a model with a covariate crossed with itself making a squared term
`NLOGISTIC Y = A X(A);` `FACTORS A.`	Fits a model with a covariate nested within a factor

The model for logistic regression is a generalization of the model used in Minitab's general linear model (GLM). Any model that is fit by the GLM command can also be fit by the logistic regression commands. For a discussion of models in general, go to How to specify the model in ANOVA on page 1252. In the logistic regression commands, Minitab assumes any variable in the model is a covariate unless the variable is specified as a factor. The FACTORS subcommand specifies which predictors are factors. In GLM, Minitab assumes any variable in the model is a factor unless the variable is specified as a covariate.

The models for Minitab's generalized linear models and predictive analytics work differently. You indicate the type of all of the predictors with the appropriate subcommand, CONTINUOUS or CATEGORICAL.

How to specify the model for factorial designs

The model for factorial designs consists of three parts: the response, factors, and the terms. You specify the response with RESPONSE. Then, you specify up to 9 factors with FACTORS. Finally, you specify the terms with TERMS. Terms consists of factors and the 2-way interactions that you want to include in the model.

Example of specifying a full quadratic model that has three factors

```
FFAC;
  Response 'BeanYield';
  Factors Nitrogen PhosAcid Potash;
  Terms C5 C6 C7 C5*C5 C6*C6 C7*C7 C5*C6 C5*C7 C6*C7;
```

The method is the same as the one used for response surface designs, but abbreviations for response surface models add square terms, such as C5*C5. Following are some examples.

Example of a main effects model that has four factors

TERMS C5 C6 C7 C8

Example of a model that has three factors with all two-way interactions

TERMS C5 C6 C7 C5*C6 C5*C7 C6*C7

Example of a model that has six factors with all two-way interactions that involve factor C5

TERMS C5 C6 C7 C8 C9 C10 C5*C6 C5*C7 C5*C8 C5*C9 C5*C10

Terms in the model are listed, separated by one or more blanks. Use an * (asterisk) to indicate an interactions term. You can omit the quotation marks around variable names if the names start with a letter and contain only letters and numbers. Otherwise, enclose the name in single quotation marks or refer to the column by number, such as C1 or C2.

A list of factors and interactions can be abbreviated. Enclose the list in parentheses and follow the parentheses by the number 2. For example, the following pair of commands is equivalent:

TERMS (A B C D)2

```
TERMS A B C D A*B A*C A*D B*C B*D C*D
```

You can use a dash, as in the rest of Minitab, to indicate a list of factors. You can also use a dash to subtract a term.

Note These two uses of a dash do not conflict.

Following are some examples:

```
TERMS C4-C7

TERMS C4 C5 C6 C7

TERMS (A B C D)2 - A*B - B*C

TERMS A B C D A*C A*D B*D C*D
```

These techniques may be combined. For example:

```
TERMS (A-C)2 - A*B D A*D

TERMS A B C A*C B*C D A*D
```

When you specify a model that has more terms than your data will support, Minitab determines which squares and interactions can be fit and fits just those.

How to specify the model for GLM

The model for GLM on page 228 is a generalization of the model used in ANOVA. Any model fit by ANOVA on page 225 can also be fit by GLM. For more information, go to How to specify the model in ANOVA on page 1252. In addition, GLM allows continuous terms that are crossed with each other and with factors, and continuous terms nested within factors. Here are some examples of these models. A is a factor and X is a continuous term.

Command language	Description

`GLM;`
`  RESPONSE Y;`
`  CONTINUOUS X;`
`  CATEGORICAL A;`
`  TERMS A X A*X.`
	Model to test homogeneity of slopes (covariate crossed with a factor)

`GLM;`
`  RESPONSE Y;`
`  CONTINUOUS X;`
`  CATEGORICAL A;`
`  TERMS A X A*X.`
	Same as previous model

`GLM;`
`  RESPONSE Y;`
`  CONTINUOUS X;`
`  CATEGORICAL A;`
`  TERMS A X X*X.`
	Quadratic model in the covariate (covariate crossed with itself)

`GLM;`
`  RESPONSE Y;`
`  CONTINUOUS X Z;`
`  CATEGORICAL A;`
`  TERMS A X Z X*X Z*Z X*Z.`
	Full quadratic in two covariates (two covariates crossed)

`GLM;`
`  RESPONSE Y;`
`  CONTINUOUS X;`
	Separate slopes for each level of A (covariate nested within a factor)

Command language	Description
CATEGORICAL A; TERMS A X; NESTED X(A).	

GLM models have some restrictions. For more information, go to Restrictions on GLM models on page 1262.

How to specify the model for response surface designs

The models for response surface designs consist of three parts: the response, the factors, and the terms. You specify the response with RESPONSE. Then, you specify up to 9 factors with FACTORS. Finally, you specify the terms with TERMS. Terms consists of factors and the two-way interactions that you want to include in the model. For example, the following specifies a full quadratic model that contains three factors.

```
RSREG;
 Response 'BeanYield';
 Factors Nitrogen PhosAcid Potash;
 Terms C5 C6 C7 C5*C5 C6*C6 C7*C7 C5*C6 C5*C7 C6*C7;
```

The method is the same as the one used in factorial designs, but models for factorial designs cannot include squares. Here are some examples:

Four factor main effects model

```
TERMS C5 C6 C7 C8
```

Three factors with all squares and all two-way interactions

```
TERMS C5 C6 C7 C5*C5 C6*C6 C7*C7 C5*C6 C5*C7 C6*C7
```

Six factors with all two interactions that involve factor C5

```
TERMS C5 C6 C7 C8 C9 C10 C5*C5 C5*C6 C5*C7 C5*C8 C5*C9 C5*C10
```

Terms in the model are listed, separated by one or more blanks. Use an * to indicate an interactions term. You may omit the quotation marks around variable names if the names start with a letter and contain only letters and numbers. Otherwise, enclose the name in single quotes or refer to the column by number such as C1 or C2.

A list of factors and interactions can be abbreviated. Enclose the list in parentheses and follow the parentheses by the number 2. For example, the following pair of commands is equivalent

```
TERMS (A B C D)2

TERMS A B C D A*A B*B C*C D*D A*B A*C A*D B*C B*D C*D
```

You can use a dash, as in the rest of Minitab, to indicate a list of factors. You can also use a dash to subtract a term.

Note These two uses of a dash do not conflict.

The following are some examples:

```
TERMS C4-C7

TERMS C4 C5 C6 C7

TERMS (A B C D)2 - A*A - A*B - B*C

TERMS A B C D B*B C*C D*D A*C A*D B*D C*D
```

These techniques may be combined. For example:

```
TERMS (A-C)2 - A*B D A*D
```

```
TERMS A B C A*A B*B C*C A*C B*C D A*D
```

When you specify a model that has more terms than your data will support, Minitab determines which squares and interactions can be fit and fits just those.

How to specify the model in ANOVA

ANOVA uses a simplified version of a model as it appears in many textbooks. The following table shows some examples.

Type of model	Minitab	Books
Two factors crossed	ANOVA Y = A B A*B	$y_{ijk} = mu + a_i + b_j + ab_{ij} + e_{k(ij)}$
Three factors crossed	ANOVA Y = A B C A*B A*C B*C A*B*C	$y_{ijkl} = mu + a_i + b_j + c_k + ab_{ij} + ac_{ik} + bc_{jk} + abc_{ijk} + el_{(ijk)}$
Three factors nested (B nested within A, and C nested within A and B)	ANOVA Y = A B(A) C(AB)	$y_{ijkl} = mu + a_i + b_{j(i)} + c_{k(ij)} + e_{l(ijk)}$
Crossed and nested design (B nested within A, and both crossed with C)	ANOVA Y = A B(A) C A*C B*C(A)	$y_{ijkl} = mu + a_i + b_{j(i)} + c_k + ac_{ik} + bc_{jk(i)} + e_{l(ijk)}$

Notice that in Minitab's models, we omit the subscripts, mu, e, and +'s that appear in textbook models. An * is used for an interaction term. Parentheses are used for nesting. For example, when B is nested within A, we write B (A), and when C is nested within both A and B, we write C (A B). Terms in parentheses are always factors in the model and are listed with blanks between them. Thus, D*F (A B E) is correct but D*F (A*B E) and D (A*B*C) are not. Also, parentheses are not used inside parentheses. Thus, C (A B) is correct but C (A B (A)) is not.

Several special rules apply only to the ANOVA command. You may omit the quotes around variable names. Because of this, variable names used in ANOVA must start with a letter and contain only letters and numbers. Alternatively, you can use C1, C2, etc., to denote data columns. You can use special symbols in a variable name, but then you must enclose the name in single quotes, as on other Minitab commands. You may not put any extra text on the ANOVA line, except after the symbol #.

You can fit reduced models. For example, Y = A B C A*B is a three-factor model with just one two-way interaction. Models, however, must be hierarchical. For example, if the term A*B*C is in the model then the terms A B C A*B A*C B*C must also be in the model. And if B(A) is in the model, then A must be also.

Because models can be quite long and tedious to type, two shortcuts have been provided. A vertical bar (or an exclamation point) indicates crossed factors, and a minus sign removes terms.

Each pair of commands below is equivalent.

ANOVA Y=A|B|C

ANOVA Y=A B C A*B A*C B*C A*B*C

ANOVA Y=A|B|C - A*B*C

ANOVA Y=A B C A*B A*C B*C

ANOVA Y=A B|C E

ANOVA Y=A B C B*C E

ANOVA Y=A|B|C|D - A*B*C - A*B*C*D

ANOVA Y=A B C D A*B A*C A*D B*C B*D C*D A*B*D A*C*D B*C*D

ANOVA Y=A|B(A)|C

ANOVA Y=A B(A) C A*C B*C(A)

In general, all crossings are done for factors separated by bars unless the cross results in an illegal term. For example, in the fifth example, the potential term A*B(A) is illegal and Minitab automatically omits it. Note, if a factor is nested you must indicate this when using the bar, as in the fifth example with the term B(A).

Several response variables can be included with one model. For example, the following commands perform three separate analyses, one for Y1, one for Y2, and one for Y3. The residuals for the first response variable are stored in C21; for the second, in C22; and for the third, in C23.

```
ANOVA Y1 Y2 Y3=A B;
 RESIDUALS C21-C23.
```

How to specify the model in MGAGE

- Use * to indicate an interaction term. For example, A*B is the interaction of the factors A and B.
- Use ( ) to indicate nesting. When B is nested within A, type B(A). When C is nested within both A and B, type C(A B). Terms in parentheses are always factors in the model and are listed with blanks between them.
- Use | or ! to abbreviate a model to indicate crossed factors.
- Use – to remove terms.

Examples

Two factors crossed: A B A*B

Three factors crossed: A B C A*B A*C B*C A*B*C

Three factors nested: A B(A) C(A B)

Crossed and nested (B nested within A, and both crossed with C): A B(A) C A*C B*C(A)

Rules

You must specify a hierarchical model.

- Crossed factors: If an interaction term is included, all lower order interactions and main effects that comprise the interaction term must appear in the model. For example, if the A and B interaction is included, denoted as A*B, then the A and B main effects terms must be included as well.
- Nested factors: If a nested factor is included in the model then the factor(s) in which they are nested must be included. If a factor is nested within multiple factors, then the interaction terms for the parent factors must be included. You cannot include interaction terms for nested factors and the factors they are nested within. For example, if C is nested in A and B, then A, B, and A*B must be included and A*C and B*C cannot be included.

When a term contains both crossing and nesting, put the * first, as in C*B(A), not B(A)*C

Terms in parentheses are always factors in the model and are listed with blanks between them. Thus, D*F (A B E) is correct but D*F (A*B E) and D (A*B*C) are not. Also, parentheses are not used inside parentheses. Thus, C (A B) is correct but C (A B (A)) is not.

Several special rules apply to the MGAGE command. You may omit the quotes around variable names. Because of this, variable names used in MGAGE must start with a letter and contain only letters and numbers. Alternatively, you can use C1, C2, etc., to denote data columns. You can use special symbols in a variable name, but then you must enclose the name in single quotes, as for other Minitab commands.

Shortcuts

Because models can be quite long and tedious to type, two shortcuts have been provided. A vertical bar (or an exclamation point) indicates crossed factors, and a minus sign removes terms.

Each pair of commands below is equivalent.

MGAGE Y=A|B|C

MGAGE Y=A B C A*B A*C B*C A*B*C

MGAGE Y=A|B|C - A*B*C

MGAGE Y=A B C A*B A*C B*C

MGAGE Y=A B|C E

MGAGE Y=A B C B*C E

MGAGE Y=A|B|C|D - A*B*C - A*B*C*D

MGAGE Y=A B C D A*B A*C A*D B*C B*D C*D A*B*D A*C*D B*C*D

MGAGE Y=A|B(A)|C

MGAGE Y=A B(A) C A*C B*C(A)

In general, all crossings are performed for factors that are separated by bars unless the cross results in an illegal term. For example, in the fifth example, the potential term A*B(A) is illegal and Minitab automatically omits it. Note, if a factor is nested you must indicate this when using the bar, as in the fifth example with the term B(A).

Missing values in exponential smoothing

Because of the nature of the computations, you cannot use the exponential smoothing commands on a time series column containing missing values (*). So what can you do if there are missing observations? Here is a possible solution:

If you have seasonal data, execute DECOMP on page 942 on the series. Replace the missing values in the series with the corresponding fitted values computed by DECOMP.

If you do not have seasonal data, replace each missing value with the mean of one or two observations immediately before and after the missing value. Another easy way to replace missing observations for non-seasonal data is to execute MA on page 952 on the series and replace the missing value with the fitted value computed by MA.

Missing values in factorial, response surface, and mixture designs

When Minitab fits factorial, response surface, and mixture models, if any column contains missing data, Minitab excludes the entire observation (row) from all calculations.

If you enter more than one response variable, Minitab handles missing data in a multivariate way. When the value of one response variable is missing for an observation (row), Minitab excludes that row in calculations for all the response variables, and displays a message. You might want to repeat the analysis separately for each response variable.

Notes on subcommands that store descriptive statistics (STATS command)

Include as many of the subcommands that store statistics as you wish, depending on which statistics you want to compute. Most often, you will use these subcommands to compute statistics for each group included in the BY columns.

The number of columns you list with any of these subcommands must equal the number of columns listed on the main STATS: Session command for storing descriptive statistics on page 144 command line, with one exception: If you omit columns from the STATS command line and use the BY subcommand, then you may store these statistics: N, NMISS, COUNT, CUMN, PERCENT, and CUMPERCENT. In this case, each subcommand must list one (and only one) column for storage.

Numbers for colors to use in session commands

Minitab provides the colors shown below. The subcommand argument is the number for the color. The color names correspond to the color names in the dialog box.

Black 16	Dark Red 24	Dark Brown 32	Chestnut 40	Dark Yellow 48	Very Dark Green 56	Indigo 64	Dark Blue 72	Dark Wine 80
Slate 17	Red 25	Ginger 33	Tangerine 41	Olive 49	Forest 57	Teal 65	Eggplant 73	Dark Magenta 81
Stone 18	Rose 114	Mustard 34	Orange 42	Chlorophyll 50	Spinach 58	Cerulean 66	Blue 74	Burgundy 82
Dark Gray 19	Terra Cotta 26	Copper 116	Gold 43	Lime 51	Dark Green 59	Azure 67	Royal Purple 75	Violet 83
Medium Gray 20	Brick 27	Wheat 35	Yellow 44	Chartreuse 52	Moss 60	Turquoise 68	Medium Blue 126	Purple 84
Gray 21	Wine 28	Peach 36	Saffron 118	Green 120	Sage 61	Cyan 69	Dark Azure 76	Grape 85
Light Gray 112	Salmon 29	Coral 37	Lemon 45	Apple 53	Dark Cyan 62	Aqua 124	Lilac 77	Plum 128
Very Light Gray 22	Pink 30	Tan 38	Light Yellow 46	Light Green 54	Sea Green 63	Light Cyan 70	Dark Sky Blue 127	Magenta 86
Pale Gray 113	Light Pink 31	Light Beige 39	Light Orange 47	Very Light Green 121	Willow 122	Light Blue 71	Light Sky Blue 78	Neon Pink 87
White 23	Pale Pink 115	Pale Beige 117	Pale Yellow 119	Pale Green 55	Spring Green 123	Pale Blue 125	Lavender 79	Dark Pink 129

Numbers for fill types to use in session commands

Minitab provides the following fill types for graph spaces (legend, area, region, bar, and so on). The subcommand argument is the number for the fill type.

Number	Fill type	Description
0		None (transparent)
1		Solid background
2		Right hatch foreground
3		Left hatch foreground
4		Right and left hatch foreground
5		Horizontal hatch foreground
6		Vertical hatch foreground
7		Horizontal and vertical hatch foreground
8		Dots foreground
9		Squares foreground

Numbers for line types to use in session commands

Minitab provides the following line types for graph objects. The subcommand argument is the number for the line type.

Number	Line type	Description
0		None
1		Solid
2		Dashes
3		Dots
4		Dash 1-Dot
5		Dash 2-Dot
6		Dash 3-Dot
7		Long dashes

Numbers to use for symbols and markers in session commands

Minitab provides the following symbols and markers for graphs. The subcommand argument is the number that corresponds to the symbol/marker type.

Number	Symbol	Description
1		Circle
2		Plus
3		Cross
4		Asterisk
5		Dot
6		Solid circle
7		Dot circle
8		Plus circle
9		Cross circle
10		Circle circle
11		Square
12		Solid square
13		Dot square
14		Cross square
15		Diamond
16		Solid diamond
17		Dot diamond
18		Plus diamond
19		Triangle
20		Solid triangle
21		Dot triangle
22		Triangle right
23		Solid triangle right
24		Dot triangle right
25		Triangle left

Number	Symbol	Description
26	◀	Solid triangle left
27	◁	Dot triangle left
28	▽	Triangle down
29	▼	Solid triangle down
30	▽	Dot triangle down

Overview of DDE session commands

The DDE dialog boxes do not generate command language as other Minitab dialog boxes do. If you want to perform DDE actions from the Command Line pane or in a macro, use the DDE session commands below.

Purpose	DDE session commands to use	Type of macros
Add a link	XDADD on page 58	Global macros and Execs only
Remove a link	XDREMOVE on page 62	Global macros and Execs only
Change a link	XDREMOVE on page 62, then XDADD on page 58 with different subcommands	Global macros and Execs only
Activate a link	XDACTIVATE on page 58	Global macros and Execs only
Deactivate a link	XDDEACTIVATE on page 60	Global macros and Execs only
Perform a one-time data transfer	XDGET on page 61	All macros
Execute an internal command	XDEXEC on page 61	All macros

Pearson residuals

The Pearson residuals are elements of the Pearson chi-square, that can be used to detect ill-fitted factor/covariate patterns. The Pearson residual defined for the jth factor/covariate pattern is given by:

$$r_j = \frac{(y_j - m_j\hat{\pi}_j)}{\sqrt{m_j\hat{\pi}_j(1-\hat{\pi}_j)}}$$

Notation

y_j	The number of successes for the j^{th} factor/covariate pattern
m_j	The number of trials for the jth factor/covariate pattern
$\hat{\pi}_j$	The estimated probability for the j^{th} factor/covariate pattern.

Plackett-Burman designs

In all cases except n = 28, you can specify the design by providing only the first column.

In the table below, we give this first column (written as a row to save space). This column is permuted cyclically to get an (n – 1) x (n – 1) matrix. Then, a last row of all minus signs is added.

For n = 28, we start with the first 9 rows. These rows are then divided into 3 blocks of 9 columns each. Then the 3 blocks are permuted (row-wise) cyclically and a last column of all minus signs is added to get the full design.

Each design can have up to k = (n – 1) factors. If you specify a K that is less than (n – 1), just the first k columns are used.

8 Runs

+ + + - + - -

13 Runs

+ + - + + + - - - + -

16 Runs

+ + + + - + - + + - - + - - -

20 Runs

+ + - - + + + + - + - + - - - - + + -

24 Runs

+ + + + + - + - + + - - + + - - + - + - - - -

28 Runs

+ - + + + + - - - - + - - - + - - + + + - + - + + - +

+ + - + + + - - - - - + + - - + - - - + + + + - + + -

- + + + + + - - - + - - - + - - + - + - + - + + - + + +

- - - + - + + + + - - + - + - - - + + - + + + - + - +

- - - + + - + + + + - - - - + + - - + + - - + + + + -

- - - - + + + + + - + - + - - - + + - - + + + - + - + +

+ + + - - - + - + - - + - - + - + - + - + + - + + + -

+ + + - - - + + - + - - + - - - - + + + - + + - - + +

+ + + - - - - + + - + - - + - + - - - + + - + + + - +

- - - - + - + - + + + - + + - - - + + + + + - - + + - + - - - +

- + - + + + - - - + + + + + - + + + - - + - - - - + - + - + + - - + -

+ + - - + + + + - + - + - - - - + + - + + + - - + + + + - + - + - - - - + + -

```
+ + - - + - + - - + + + - + + + + + + - - - + - + + + - - - - - + - - - + + - + - + + -
```

48 Runs

```
+ + + + + - + + + + - - + - + - + + + - - + - - + + - + + - + - + + - - - + - + - + + - - - - + - - - -
```

Prompting a user for information

READ, SET, and TSET have a special feature that allows you to ask users questions and then use their answers in the macro. A macro will pause for user input if you use READ, SET, or TSET with the subcommand FILE with the special file name TERMINAL. TERMINAL tells Minitab to wait for input from the keyboard. READ, SET, and TSET also have other subcommands.

Reading in data from a text file

Most applications, such as spreadsheets and data bases, can save data in an ASCII file, also called a data or text file. You can use READ (for information, go to READ data into a matrix on page 44 or READ data into columns on page 42) or SET on page 81 to import this type of file into Minitab. Minitab assumes the default file extension DAT for use with data files. You do not need to type DAT as part of the file name within single quotation marks if the file has a DAT extension. For example, the following commands import the file named MYDATA.DAT into C1 through C10. MYDATA contains 10 columns of numbers, separated by spaces or commas.

```
READ C1-C10;
 FILE "MYDATA".
```

If the file is not in your default directory, include the path information as part of the file name within single quotation marks. If the file extension is not DAT, include the file extension as well.

If the file includes text columns, if the columns are separated by tabs, or if missing values are designated by a blank, you must use additional subcommands.

Note The menu command **File** > **Open Worksheet** also opens ASCII (text) files, and many other types of files as well. The menu command provides options that are not available with READ or SET.

Generalized linear model diagnostics and residual analysis

Following any modeling procedure, you should assess the validity of your model. The generalized linear model (GZLM) command includes a collection of subcommands that store residuals and diagnostic statistics. These residuals and diagnostic statistics allow you to do the following:

- identify factor/covariate patterns which are poorly fit by the model
- identify factor/covariate patterns which have a strong influence on the estimated parameters
- identify factor/covariate patterns which have a large leverage

Hosmer and Lemeshow suggest that you interpret these diagnostics jointly to understand any potential problems with the model.

Subcommand	What the subcommand stores	Use the stored statistics to identify
RESIDUALS	Pearson or deviance residuals	Poorly fit factor/covariate patterns

Subcommand	What the subcommand stores	Use the stored statistics to identify
SRESIDUALS	Standardized Pearson residuals or standardized deviance residuals	Poorly fit factor/covariate patterns
TRESIDUALS	Deleted Pearson residuals or deleted deviance residuals	Poorly fit factor/covariate patterns
DCHISQUARE	Changes in the Pearson chi-square when the j^{th} factor/covariate pattern is removed	Poorly fit factor/covariate patterns
DDEVIANCE	Changes in the deviance when the j^{th} factor/covariate pattern is removed	Poorly fit factor/covariate patterns
DSBETA	Changes in the coefficients when the j^{th} factor/covariate pattern is removed, based on standardized Pearson residuals	Factor/covariate patterns with a strong influence on parameter estimates
DBETA	Changes in the coefficients when the j^{th} factor/covariate pattern is removed, based on Pearson residuals	Factor/covariate patterns with a strong influence on parameter estimates
HI	Leverages of the j^{th} factor/covariate pattern	Factor/covariate patterns with a large leverage

The analyses store deviance residuals unless you use the subcommand PEARSON. The GZLM command has a collection of residual plots. On these plots, you can use Minitab's graph brushing capabilities to identify poorly fit and influential factor/covariate patterns.

Residual analysis and regression diagnostics

Further analysis usually follows the fitting of any type of model, such as regression. The most important part is the plotting of residuals. For example, many analyses rely on the assumption that residuals are normally distributed and are constant, that is, the residuals do not exhibit a trend as a function of the response variable. Minitab provides subcommands to graphically illustrate whether or not these assumptions have been met.

What you want to plot	Graph subcommand to use
Histogram of the residuals	GHISTOGRAM
Normal plot of the residuals	GNORMAL
Residuals versus fits ($\hat{Y}$)	GFITS
Residuals versus order	GORDER
A layout of a histogram of the residuals, normal plot of the residuals, residuals versus fits, and residuals versus order residuals versus predictor variables	GFOURPACK
Residuals versus predictor variables	GVARIABLE

You can specify the type of residual to plot-raw residuals, standardized residuals, or Studentized deleted residuals-using the RTYPE subcommand. You can also store these residuals using the subcommands RESIDUALS, SRESIDUALS, and TRESIDUALS, respectively.

In addition, REGRESS has a collection of subcommands (HI, RESIDUALS, TRESIDUALS, COOKD, DFITS) that store diagnostics to help you find unusual observations. You should check unusual observations to make sure they are correct. If correct, you should try to understand why they are unusual and what effect they have on the regression equation.

Restrictions on GLM models

Although models can be unbalanced in GLM on page 228, they must be "full rank". Thus, there must be enough data to estimate all the terms in your model. For example, suppose you have a two-factor crossed model with one empty cell. Then you can fit the model GLM Y = A B, but not GLM Y = A B A*B. Don't worry about figuring out whether or not your model is of full rank. Minitab will tell you if it is not. In most cases, eliminating some of the high order interactions in your model (assuming, of course, they are not important) will solve your problem.

There is a restriction on the subscripts used for nested terms as nesting must be balanced. Suppose A has 3 levels, and B is nested within A. If B has 4 levels within the first level of A, B must have 4 levels within the second and third levels of A. Minitab will tell you if you have unbalanced nesting. In addition, the subscripts used to indicate the 4 levels of B within each level of A must be the same. Thus, the four levels of B cannot be (1 2 3 4) in level 1 of A, (5 6 7 8) in level 2 of A, and (9 10 11 12) in level 3 of A.

Session commands that are not allowed in macros

Commands and subcommands that open different projects or restart Minitab are not allowed in macros. In addition, local macros cannot include commands that do the following:

- Make different worksheets active.
- Require Minitab to pick storage locations after the last column in use.

Where subcommands are listed below, the command itself is permitted. For example, you can use the COPY command in local macros, if you do not use the NEWWS, AFTER, or STORE subcommands.

Not allowed in local macros

- COPY
  - NEWWS
  - AFTER
  - STORE
- EXECUTE
- NEW
- RETRIEVE
- SAVE
- SORT
  - NEWWS
- SPLIT
- STACK
  - NEWWS
- SUBSET
- TRANSPOSE
  - NEWWS
  - AFTER

- UNSTACK
  - NEWWS
  - AFTER
- WOPEN
- WORKSHEET

Not allowed in global macros

- NEW
  - PROJECT
- RESTART
- RETRIEVE
  - PROJECT

Simplex lattice design descriptions

Degree of design (m)	Number of components (q)
1	2 to 10
2	2 to 20
3	2 to 17
4	2 to 11
5	2 to 8
6	2 to 7
7	2 to 6
8	2 to 5
9	2 to 5
10	2 to 5

Time unit subcommands

Use these subcommands to specify when consecutive observations in a given variable were made. Use START and INCREMENT to specify the starting point of the first observation and the frequency of consecutive points, all in terms of the given unit. For example, a series of observations beginning in March and taken every other month thereafter would be represented by the following commands.

```
MONTH;
START 3;
INCREMENT 2.
```

These subcommands specify the rate at which data were collected, not the number and frequency of ticks on the resulting graph. For information on tick placement and custom labels, go to SCALE on page 1149.

You can specify up to three time unit subcommands in one of the following combinations.

Time units	Supported combinations
Clock	DAY; HOUR DAY; HOUR; MINUTE HOUR; MINUTE HOUR; MINUTE; SECOND MINUTE; SECOND
Calendar	DAY; MONTH DAY; MONTH; YEAR MONTH; YEAR MONTH; QUARTER MONTH; QUARTER; YEAR QUARTER; YEAR

A scale level is shown for each time unit. You can specify a start value for each.

INDEX
Specifies that consecutive integers be used on the time scale. The default START value is 1, and all integers are valid START arguments.

STAMP C [C...C]
Specify up to three columns that contain the time values for the observations.

SECOND
Specifies that observations were taken in units of seconds. The default START value is 0, and 0 through 59 are valid START arguments.

MINUTE
Specifies that observations were taken in units of minutes. The default START value is 0, and 0 through 59 are valid START arguments.

HOUR
Specifies that observations were taken in units of hours. The default START value is 0, and 0 through 23 are valid START arguments.

DAY
Specifies that observations were taken in units of days. The default START value is 1, and 1 through 31 are valid START arguments, depending on the month.

MONTH
Specifies that observations were taken in units of months. The default START value is 1, and 1 through 12 are valid START arguments.

QUARTER
Specifies that observations were taken in units of quarters. The default START value is 1, and 1 through 4 are valid START arguments.

YEAR
Specifies that observations were taken in units of years. The default START value is 1900, and all positive integers from 1000 to 4000 are valid START arguments.

START K [K...K]

START C

Specifies the starting point of the given unit or units. START takes as many arguments as there are time unit subcommands. The Start values should be applied from outermost to innermost scale levels, the opposite of calendar combinations. You may also enter a column containing the start values. If you have multiple graph variables you can enter one set of START values for all, or one for each.

INCREMENT K

Specifies the time scale increment. INCREMENT always refers to the smallest specified time unit subcommand and its argument is in terms of this unit. An argument of 1 assigns successive observations to consecutive seconds, minutes, hours, and so on. An argument of 2 assigns successive observations to every other second, minute, hour, and so on.

Using groups in graphs

You define groups by specifying one, two, or three group variables with data display subcommands in graphs that can use groups with data display subcommands. For more information, go to Graphs that use groups with the data display subcommands on page 1241.

The distinct values of a single variable or the combinations of distinct values of several variables define the groups. The following table below defines groups in five different ways. For numeric variables, such as Quarter (Qtr) or Year (Yr), groups are assigned in increasing order of the values of the group variable. For text variables, such as AdAgency, groups are defined in the order of appearance in the data set. The first value of AdAgency is Omega, so it is group 1. The second distinct value to appear is Alpha, so it is group 2.

C1	C2	C3	C4	C5	C6	C7	C8	C9
Row	Qtr	Yr	AdAgency	Groups Qtr	Groups Yr	Groups AdAgency	Groups Y by AdA	Groups AdA by Y
1	1	2015	Omega	1	1	1	1	1
2	2	2015	Omega	2	1	1	1	1
3	3	2015	Alpha	3	1	2	2	3
4	4	2015	Alpha	4	1	2	2	3
5	1	2016	Alpha	1	2	2	4	4
6	2	2016	Alpha	2	2	2	4	4
7	3	2016	Omega	3	2	1	3	2
8	4	2016	Omega	4	2	1	3	2

When you use several group variables, the values of the first variable vary the slowest, the second vary the second slowest, and the last vary the fastest. This means that specifying group variables in a different order can affect which group comes first, second, and so on.

For example, look at the fourth group definition, Year by AdAgency. The first group is the first Year and first AdAgency; the second group is still the first Year but now the second AdAgency. The third group is the second Year and first AdAgency; the fourth group is the second Year and second AdAgency. The last group definition, AdAgency by Year, uses the same variables but in the opposite order. Notice how the group assignments change.

For more information on groups, go to Assigning attributes to groups on page 1215 and Assigning attributes with multiple graphs and groups on page 1215.

Using READ with FORMAT

- When you use READ and have typed all your data, type END on page 39 in the first three spaces of the next line. For more information, go to READ data into a matrix on page 44 or READ data into columns on page 42.
- Numbers that are read under F format must be typed in the spaces indicated by the format, with no extra text. Numbers should normally be typed right-justified (that is, with no blanks following the number in the field).
- Currency values that are read under C format must be typed in the spaces indicated by the format, with no extra text other than the currency symbol. Currency values should normally be typed right-justified (that is, with no blanks following the number in the field).
- You can indicate missing data by putting a single * in the field or by leaving the field blank. A count of blank fields is given after all the data have been read.
- If there are errors or warnings, fairly extensive messages are given for the first 10 rows.
- Use a comma to separate format items.

Using READ without subcommands

- READ (for information, go to READ data into a matrix on page 44 or READ data into columns on page 42) ignores text on data lines. However, a data line cannot begin with a valid command name.
- You can type data with or without a decimal point and in exponential (scientific) notation. Thus, the following values are all valid and equivalent:

  123.4

  1.234E+2

  1.234E2

- You can use the missing data code (*) as data.
- READ ignores lines that are blank or that contain only spaces.
- For information on using READ to put numbers into a matrix, go to READ data into a matrix on page 44.
- You can type comments on a data line after the symbol #.
- You cannot enter stored constants (such as K1) in place of numbers on a data line after READ, but you can with SET on page 81.
- If a row is incomplete—that is, it contains fewer values than are listed as columns on the main READ command line—then Minitab reads all the values in the row and sets the end values in the row as missing.

Using the SYMBOL and COLOR subcommands for PLOT

The data display subcommands determine the basic appearance of a graph produced by a supported graph command.

You can specify more than one data display subcommand in a graph (for example, using SYMBOL and CONNECT together). You can also specify the same data display subcommand several times in the same graph (for example, using two SYMBOL commands to identify distinct points differently).

Many data display subcommands have subcommands of their own. This section discusses how the data display subcommands work in general.

Minitab allows you to display different subsets of points in several ways. You can display:

- Every point on the graph in the same way (Graph Mode)
- Each individual point on a graph differently (Point Mode)
- Points belonging to different groups differently (Group Mode)

These concepts are explained using one graph (PLOT), one data display subcommand (SYMBOL), one of the subcommands of SYMBOL (COLOR), and the Marketing Data. Minitab uses integers for the colors. For the examples in this section, we use the colors 1 = black, 2 = red, 3 = green, and 4 = blue.

If you use only one argument with COLOR, the argument applies to every point on the graph (this is Graph mode.). For example:

```
PLOT Sales*Advertis;
 SYMBOL;
 COLOR 4.
```

displays all points with the same symbol type and colors all the symbols the same.

If you use several arguments with COLOR, and no arguments with SYMBOL, the COLOR arguments are applied to successive points in the data set (hence, Point mode). For example:

```
PLOT Sales*Advertis;
 SYMBOL;
 COLOR 1 1 1 2 1 1 1 1.
```

colors the first 8 points the color that corresponds to 1 except the fourth, which is the color that corresponds to 2. Further points are the default color.

Group mode

Suppose we add a grouping variable, Year, which divides the data into two groups, 1 and 2 (Group mode):

```
PLOT Sales*Advertis;
 SYMBOL Year.
```

automatically gives points for each year a different symbol type and a different color.

This automatic assignment of different symbol types to different groups continues if you use COLOR, SIZE, ETYPE, ECOLOR, or ESIZE. For example:

```
PLOT Sales*Advertis;
 SYMBOL Year;
 COLOR 4.
```

gives the points for each year a different symbol type with the same color.

If we use two arguments with COLOR (or TYPE, SIZE, ETYPE, ECOLOR, or ESIZE), the first applies to the first group and the second to the second group, and so on (still Group mode). For example:

```
PLOT Sales*Advertis;
 SYMBOL Year;
 COLOR 2 3.
```

colors all symbols in first year the color that corresponds to 2 and all symbols in the second year the color that corresponds to 3.

Now let's look at multiple graphs, first without a grouping variable. If we use just one argument with COLOR, then all points in all graphs are given the same color. If we use several arguments with COLOR, then the arguments are assigned to the points. For example:

```
PLOT Sales*Year Advertis*Year;
 SYMBOL;
 COLOR 2 3.
```

colors the first point in both graphs the color that corresponds to 2 and the second point in both graphs the color that corresponds to 3.

If you have several graphs in one command and use a grouping variable with SYMBOL, then the colors are assigned to the different groups, not the different graphs. For example:

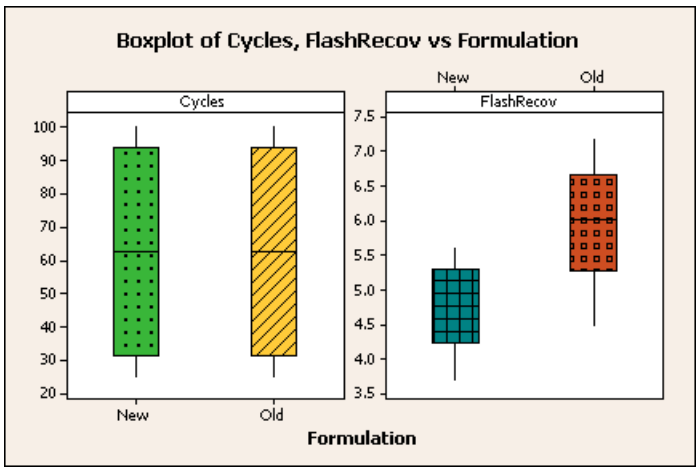
```
PLOT Sales*Year Advertis*Year;
 SYMBOL AdAgency;
 COLOR 2 3.
```

makes the first group in AdAgency the color that corresponds to 2 in both graphs. The second group in AdAgency is the color that corresponds to 3 in both graphs.

To assign different colors to different graphs when using groups, use separate PLOT commands.

Using the POSITION and MODEL subcommands

The POSITION subcommand allow you to affect graph elements (such as data labels, symbols, boxes, bars, or axis labels) for specific data values. The MODEL subcommand extends this capability to paneled graphs.



The green boxplot in the Cycles panel is POSITION 1; MODEL 1. The yellow boxplot to its right is POSITION 2; MODEL 1.

The blue boxplot in the FlashRecov panel is POSITION 1; MODEL 2. The orange boxplot to its right is POSITION 2; MODEL 2.

The following session commands create this graph.

```
BOXPLOT CYCLES * FORMULATION FLASHRECOV * FORMULATION;
  PANEL;
  IQRBOX;
    POSITION 1;
      MODEL 1;
      TYPE 8;
      COLOR 50;
    ENDPOSITION;
    POSITION 2;
      MODEL 1;
      TYPE 2;
      COLOR 43;
    ENDPOSITION;
    POSITION 1;
      MODEL 2;
      TYPE 7;
      COLOR 62;
    ENDPOSITION;
    POSITION 2;
      MODEL 2;
      TYPE 9;
```

```
    COLOR 27;
  ENDPOSITION;
  TYPE 4;
  COLOR 50;
OUTLIER;
TITLE;
  PSIZE 14;
FOOTNOTE;
  FPANEL;
NODTITLE.
```

Valid format items

Item	Format	Description
Fw	F format	For reading numbers. Fw is the basic format item. Fw reads a number from a field of width w. For example, F10 says to read the number in the next 10 spaces.
Fw.d	F format	For reading numbers with number of decimal places. The format, Fw.d, inserts a decimal point d spaces from the right of the number. Regardless of local regional settings for language, a period character must be placed before the specified number of decimal places. For example, F10.4 takes the number from the next 10 spaces on the data line, and inserts a decimal point before the last 4 digits. If there is a decimal point in the number, it overrides the d.
Aw	A format	For text (alpha) data. The A format item is for text (alpha) data. The field width, w, may be up to 80. Columns input with the A format can be output by WRITE: Session command for writing data to the screen or a data file on page 50 and PRINT: Session command for displaying columns, constants, or matrices in the output pane on page 42, with or without a FORMAT subcommand, and can be converted to numbers with CONVERT: Session command for converting text data to numeric data, and numeric data to text data on page 122.
DT	DT format	For date/time data. The DT format is a special format unique to Minitab for reading date/time data. The example DT8m/d/yy reads data from the first 8 spaces in the format m/d/yy into a date/time column. You need to provide a space value for the width only if you are reading from a file.
%w.d	% format	For percentages with numbers of decimal places. The %w.d format is for reading percentage data. Minitab formats percentage data as numeric data.
X	X format	Skips a space.
n	Repeat factor	
(	Open parenthesis	
)	Close parenthesis	

Item	Format	Description
,	Comma	Separates format items.
/	Slash	Moves to a new data line.
Iw		
Gw		
Gw.d		
Ew		
Ew.d		
Dw		
Dw.d		